VECTOR ANALYSIS

With an Introduction
to
TENSOR ANALYSIS

PRENTICE-HALL MATHEMATICS SERIES

ALBERT A. BENNETT, EDITOR

VECTOR ANALYSIS

With an Introduction to TENSOR ANALYSIS

BY

JAMES HENRY TAYLOR
PROFESSOR OF MATHEMATICS
THE GEORGE WASHINGTON UNIVERSITY

NEW YORK
PRENTICE-HALL, INC.

First printing........................October, 1939
Second printing........................June, 1945
Third printing........................July, 1946
Fourth printing........................August, 1947

Preface

THIS book is an effort to present, in as simple a manner as possible, an introduction to Vector Analysis that will naturally lead to its extension, Tensor Analysis. Any treatment meeting this objective must, it seems to me, necessarily involve an adequate discussion of the theory of linear dependence and the notion of invariants with respect to a group of transformations. The concept of a system of base vectors has been introduced early and used extensively in the belief that it contributes greatly to a clear understanding of the operations in Vector Analysis, including the operation of differentiation.

At the beginning, the notion of a vector is motivated by means of translations carried out in terms of affine coördinate systems. We then lay down, postulationally, the laws that shall govern certain operations of the vectors and their relations to points in a given affine coördinate system. In §7 and §9 we introduce the notion of a vector as an invariant with respect to the group of affine transformations, and in Chapter IV this idea is enlarged upon with respect to the more general group of analytic transformations.

Care has been taken not to make use of a specialized choice of base vectors, or coördinate system, until it is clear that such a choice will result in a simplified description or procedure concerning the quantities involved. In this way the invariantive character, with respect to a given class of coördinate systems, of the theorems and procedures receives an emphasis which is deserved, and which is frequently obscured by almost exclusively employing an **i**, **j**, **k** system of base vectors.

Applications of a rather wide and diverse nature have been considered or indicated. However, the treatment has been presented as a course in mathematics, and the applica-

tions have been chosen for their simplicity and illustrative value and not with regard to their probable importance in the field of application.

A mathematical maturity on the part of the reader, which may be expected of one who has studied a first course in the calculus, has been assumed. I also presuppose the reader's willingness to amplify, when necessary, his knowledge of certain topics by additional reading. It is believed that the references cited will provide suitable direction for such reading and point the way toward a more exhaustive discussion of certain ideas.

Many exercises are provided, some of which yield significant results and might well be presented as a portion of the text. However, when these theorems are within the capacity of the student to demonstrate, it has seemed best to present them as exercises. Generally speaking, the exercises are to be regarded as an integral part of the development.

In Chapter III, "Integral Calculus of Vectors," only an intuitive approach to the subject has been offered. The difficulties in this connection are perhaps too serious to be rigorously disposed of here.

Space is lacking for complete acknowledgments to those who have made this book possible. Mr. Gilbert A. Hunt, Jr., has rendered invaluable assistance in the preparation of the manuscript. Among the works pertaining to the subject I have found those of Veblen, Weyl, Gibbs-Wilson, and Juvet especially helpful. I am greatly indebted to my former teachers, Professor A. C. Lunn of the University of Chicago, for an understanding of the subject as a whole, and the late Professor E. H. Moore of the same institution for an appreciation of the power of a basis in this and other connections. My indebtedness to my many students of the subject is cheerfully acknowledged; the contributions they have made, though difficult to measure, are very real.

J. H. T.

Contents

CHAPTER I

Algebra of Vectors

§1. Affine Coördinate Systems and Translations

1.1 The coördinate axiom.

References:[1] *Carathéodory* (29), p. 18; *Schreier-Sperner* (49), I, p. 5; *Reynolds-Weida* (48), p. 1; *Veblen* (50), p. 13; *Osgood-Graustein* (46), p. 2.

Let O and E be two distinct points on a line l. The point O is called the *zero point*, or *origin*, of the coördinate system to be determined; the point E is called the *unit point*. The point O divides the line l into two *half-lines:*

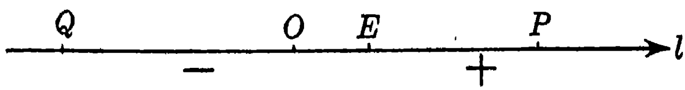

Fig. 1

that on which E lies is known as the *positive* half-line, and the other is known as the *negative* half-line. If P is any point on the positive half-line, let there be ordered to P the unique[2] real number x,

$$x = \frac{\overline{OP}}{\overline{OE}},$$

defined by the ratio of the lengths of the line segments $\overline{OP}$ and $\overline{OE}$. The number x is called the *coördinate* of P. If Q is a point on the negative half-line, its coördinate x is defined by

$$x = -\frac{\overline{OQ}}{\overline{OE}}.$$

[1] See bibliography beginning on page 174 of this book.

[2] The word "unique" as used in mathematics is equivalent to "one and only one."

We have, then, corresponding to each point P on the line l, a unique real number x; conversely, each real number x is the correspondent of a unique point P on l. A one-to-one reciprocal correspondence[3] between points of a line and numbers of a number system is called a *coördinate system*. The particular coördinate system above described is known as a Cartesian coördinate system on the line. That such a coördinate system exists is an *assumption* which we make. It is this assumption that makes the subject "Analytic Geometry" possible. We note that in the above system the coördinate of the point O is 0 (zero), and the coördinate of E is 1 (one).

We assume the parallel postulate of Euclidean geometry: Through any point not on a given line there passes one and only one line *parallel* to the given line. In order to remove the exceptional character of points on the given line, we shall regard a line as being parallel to itself.

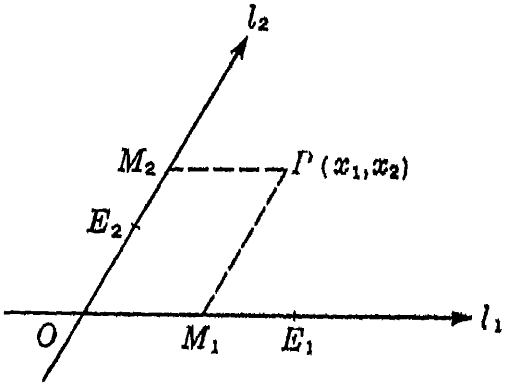

Fig. 2

Now let l_1 and l_2 be two distinct lines which meet in the point O. We establish a coördinate system for the plane containing the lines l_1 and l_2. Set up a coördinate system on each of the lines l_1 and l_2 as above, in which the common point O is taken as the zero-point on each line. Let P be

[3] Two sets of objects S and S_1 are said to be in a one-to-one reciprocal correspondence, if to each element of S there corresponds a unique element of S_1, and each element of S_1 is the correspondent of a unique element of S.

any point in the plane, and through P let there be passed lines respectively parallel to l_2 and l_1, which meet these lines in M_1, M_2 (Fig. 2). Let the coördinate of M_1 be x_1 and the coördinate of M_2 be x_2. Then the ordered pair of numbers (x_1, x_2) are called the *coördinates* of the point P. Conversely, if x_1 and x_2 are given, they determine a unique point whose coördinates they are. Thus we have a one-to-one reciprocal correspondence between points of the plane and ordered pairs of real numbers. Frequently this is called a parallel coördinate system because the locus of points in the plane satisfying the equation

$$x_1 = \text{constant, or } x_2 = \text{constant}$$

is in each case a straight line parallel to one of the coördinate axes OX_1 or OX_2. However, we shall use the term *affine* coördinate system as its description.

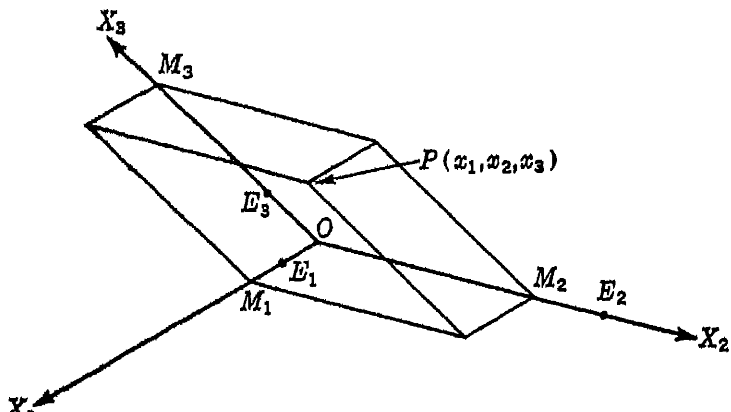

Fig. 3

The extension of the notion of an affine coördinate system to space is readily made. Let three distinct planes having a unique point O in common meet in pairs in the lines OX_1, OX_2, OX_3, and on each of these lines establish a coördinate system having the common point O as the zero-point. Let P be any point in the space, and let planes be passed through P respectively parallel to the

coördinate planes, determining the points M_1, M_2, M_3[4] (Fig. 3). Let the coördinate of M_1 on the line OX_1 be x_1. Similarly, let the coördinate of M_2 on OX_2 be x_2, and of M_3 on OX_3 be x_3. Then the numbers of the ordered set of numbers (x_1, x_2, x_3) are called the *coördinates* of the point P. Clearly the coördinates of O are (0, 0, 0), and those of E_1, E_2, E_3 are respectively (1, 0, 0), (0, 1, 0), (0, 0, 1).

A relief map is an instance of an affine coördinate system. We assume a uniform scale in the East-West direction, in the North-South direction (not necessarily the same as in the East-West direction), and for measurement of altitude. On such a map we have a means for determining whether a point is three times as far east from a given meridian as another point. Also we can compare the altitude of two given points. However, we are not in position to obtain the distance between any two arbitrarily given points, for we make no assumption at this time concerning the comparison of scales along the three directions.

1.2 Translations.

References: *Weyl* (53), p. 11; *Reynolds-Weida* (48), p. 167; *Veblen-Young* (51), II, p. 74; *Osgood-Graustein* (46), p. 330.

If a rigid body is displaced so that the path traced out by each point of the body is a straight line, the displacement is called a *linear displacement* or a *translation*. Let l be a line, and let A and A' be two distinct points on l. Consider now the *translation* of all the points of space which transforms, or carries, A into A'. The translation has the property that it not only leaves the line l invariant but it also leaves every line invariant which is parallel to l. By stating that the translation leaves the line l *invariant* is simply meant that, under the translation, the correspondent of every point on l is a point on l. That is, collectively

[4] The figures are intended as an aid in the exposition, and when the relationships of the various elements involved are clear from the figure, the detailed description of such relationships will frequently be omitted in the text.

the set of points constituting l is carried into itself. A translation will clearly be completely described by giving the point A and the corresponding point A'. It will be, however, equally well described by giving any point B together with its correspondent B'.

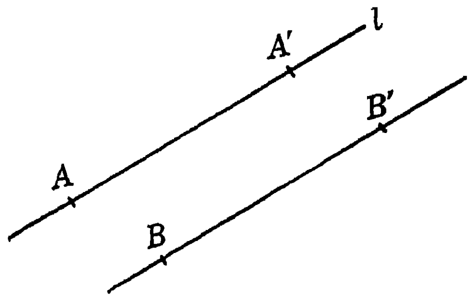

Fig. 4

A necessary and sufficient condition that two directed line segments AA' and BB' describe the same translation is that they satisfy the three properties:

(1) AA' and BB' are *parallel*.

(2) The segments AA' and BB' are *congruent* in the sense that there exists a *translation* which carries the segment AA' into BB'.

(3) The segments have the *same sense* of *direction*.[5]

Since we are for the present limiting the displacements employed to translations, we are unable to consider the congruence of two segments which are not parallel. After the introduction of a metric, this restriction will be removed. In the example of the relief map, translations alone suffice to make a comparison of the altitude of two points, or of their distance east or west of a given North-South line (meridian).

Let corresponding points of a translation A and A' have coördinates (a_1, a_2, a_3) and (a'_1, a'_2, a'_3), respectively, in the

[5] Evidently (2) implies (1); the latter is explicitly stated for emphasis.

same coördinate system, and set $\alpha_1 = a_1' - a_1, \alpha_2 = a_2' - a_2$, $\alpha_3 = a_3' - a_3$. Now if (x_1, x_2, x_3) are the coördinates of a point P, the coördinates (x_1', x_2', x_3') of the corresponding point P' under the translation are given by

$$x_1' = x_1 + \alpha_1,\ x_2' = x_2 + \alpha_2,\ x_3' = x_3 + \alpha_3.$$

That is, in terms of affine coördinates a real[6] translation is a correspondence of the form

$$\mathrm{T}:\quad \begin{aligned} x_1' &= x_1 + \alpha \\ x_2' &= x_2 + \beta \\ x_3' &= x_3 + \gamma, \end{aligned}$$

where α, β, γ are real numbers.

Conversely, any transformation of the form T can be interpreted as a translation. First, we deduce the equations of a line in space. Suppose the line is determined as passing through two distinct points $A(a_1, a_2, a_3)$ and $A'(a_1', a_2', a_3')$. Let $P(x_1, x_2, x_3)$ be an arbitrary point on the required line. Then from similar figures we have

$$\frac{x_1 - a_1}{a_1' - a_1} = \frac{x_2 - a_2}{a_2' - a_2} = \frac{x_3 - a_3}{a_3' - a_3}$$

or

$$\frac{x_1 - a_1}{\alpha} = \frac{x_2 - a_2}{\beta} = \frac{x_3 - a_3}{\gamma},$$

the equations of the line passing through A and A'. Conversely, equations of this form define a straight line when α, β, γ are not all zero.

Given, then, a transformation

$$\mathrm{T}:\quad \begin{aligned} x_1' &= x_1 + \alpha \\ x_2' &= x_2 + \beta \\ x_3' &= x_3 + \gamma \end{aligned}$$

with α, β, γ real numbers which we suppose not all zero,

[6] The term "real" is merely descriptive of the fact that we are, for convenience, restricting the numbers α, β, γ to be real numbers—that is, not complex numbers or numbers of some other number system.

consider the line defined by

$$\frac{x_1 - b_1}{\alpha} = \frac{x_2 - b_2}{\beta} = \frac{x_3 - b_3}{\gamma},$$

where (b_1, b_2, b_3) is an arbitrary point. If a point $P(x_1, x_2, x_3)$ is on this line, its correspondent $P'(x_1', x_2', x_3')$ is likewise. For x_1', x_2', x_3' satisfy the equations

$$\frac{x_1' - b_1 - \alpha}{\alpha} = \frac{x_2' - b_2 - \beta}{\beta} = \frac{x_3' - b_3 - \gamma}{\gamma},$$

and hence

$$\frac{x_1' - b_1}{\alpha} = \frac{x_2' - b_2}{\beta} = \frac{x_3' - b_3}{\gamma}.$$

That is, every line of the family of parallel lines specified by direction numbers α, β, γ is invariant under the transformation T. It is also clear that the line segments joining corresponding points satisfy the conditions (1), (2), (3) listed above. Hence equations T define a *translation*.

If T_1 and T_2 are two real translations, then there exists a real translation T_3 which is the resultant or "product" of the translations T_1 and T_2, in the order T_1 followed by T_2,

$$T_3 = T_2T_1.$$

Let

$$T_1:\ \begin{aligned} x_1' &= x_1 + \alpha_1 \\ x_2' &= x_2 + \beta_1 \\ x_3' &= x_3 + \gamma_1 \end{aligned} \qquad T_2:\ \begin{aligned} x_1'' &= x_1' + \alpha_2 \\ x_2'' &= x_2' + \beta_2 \\ x_3'' &= x_3' + \gamma_2. \end{aligned}$$

Then

$$T_3:\ \begin{aligned} x_1'' &= (x_1 + \alpha_1) + \alpha_2 \\ x_2'' &= (x_2 + \beta_1) + \beta_2 \\ x_3'' &= (x_3 + \gamma_1) + \gamma_2 \end{aligned}$$

or

$$T_3:\ \begin{aligned} x_1'' &= x_1 + (\alpha_1 + \alpha_2) \\ x_2'' &= x_2 + (\beta_1 + \beta_2) \\ x_3'' &= x_3 + (\gamma_1 + \gamma_2) \end{aligned}$$

or

$$T_3:\ \begin{aligned} x_1'' &= x_1 + \alpha_3 \\ x_2'' &= x_2 + \beta_3 \\ x_3'' &= x_3 + \gamma_3 \end{aligned}$$

where α_3, β_3, γ_3 are real numbers, since each is the sum of two real numbers. That is, T_3 is a real translation. Moreover, since we assume addition to be commutative for the real numbers, it follows that the resultant of two real translations is likewise commutative; that is, $T_3 = T_2T_1 = T_1T_2$. This theorem has the obvious geometric interpretation or graphical representation indicated in Fig. 5.

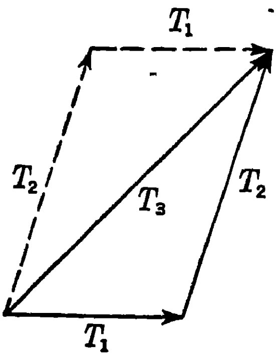

Fig. 5

1.3 Group.

References: *Veblen-Young* (51), I, p. 66; *Bôcher* (27), p. 80; *Young* (55), p. 89; *Schreier-Sperner* (49), II, p. 7.

One of the most important concepts in mathematics is that of a group.

Definition: A class G of elements which we denote by $a, b, c, \ldots$ is said to form a *group* with respect to an *operation* or *law of combination* $\bigcirc$ acting on pairs of elements of G, if the following postulates are satisfied:

(1) For every pair of elements a, b (equal or distinct) of G, the result $a \bigcirc b$ of acting with the operation $\bigcirc$ on the elements in the order given is a uniquely determined element of G. (Closure property.)

(2) $(a \bigcirc b) \bigcirc c = a \bigcirc (b \bigcirc c)$. (Associative property.)

(3) There exists in G an element i such that $a \bigcirc i = a$ for every element a of G. (Right identity element.)

(4) For every element a in G there exists an element a' in G such that $a \bigcirc a' = i$. (Right inverse.) The group is said to be *commutative* if the additional postulate is satisfied:

(5) $a \bigcirc b = b \bigcirc a$ for every pair of elements a, b in G.

Instances of a group are (1) the set of integers, including zero, with respect to addition; (2) the set of real numbers

with respect to addition; (3) the set of all real numbers except zero with respect to multiplication.

Exercises

1.1. Obtain the translation which carries the point (1, 2, 3) into (−2, 3, −1).

1.2. The points $A\,(1, 4)$, $B\,(2, -3)$, $C\,(-3, 5)$, $D\,(-4, 12)$ are the vertices of a parallelogram.

1.3. Prove directly from the equations of transformation that the set of real translations in space constitutes a commutative group with respect to the operation of forming their resultant.

1.4. Let x, y be rectangular Cartesian coördinates in the plane. The transformation

$$\begin{aligned} x &= x' \cos\theta - y' \sin\theta \\ y &= x' \sin\theta + y' \cos\theta \end{aligned}$$

is called a *rotation about the origin.* Show that

(1) The set of real rotations about the origin constitutes a commutative group with respect to the operation of forming their resultant.

(2) Any circle with center at the origin is invariant with respect to the group of rotations about the origin.

1.5. The three cube roots of unity constitute a group with respect to multiplication. Interpret geometrically.

1.6. Prove the following theorems concerning a group:

(1) For any two elements a and b of G there is a unique element c of G such that $c \bigcirc a = b$.

(2) Given any element a of G, there exists a unique left inverse; that is, an element a^* such that $a^* \bigcirc a = i$, i being an identity element postulated above.

(3) $i \bigcirc a = a$ for every element a in G.

(4) There is only one identity element in G.

(5) For each element in G there is only one inverse.

(6) The inverse of a product is the product of the inverses in the reverse order.

(7) Forming the inverse of an element of G is an operation of period 2. (An operation is said to be of period 2 if, when applied twice to any element of a class, it yields that element. For instance, taking the negative of a number is such an operation.)

§2. Scalars and Vectors

References: *Gibbs-Wilson* (7), pp. 1–4; *Ames-Murnaghan* (24), pp. 1–6.

In our work we shall need to distinguish carefully two kinds of quantities, *scalar* quantities and *vector* quantities.

A scalar quantity is a quantity whose measure can be described by a single number. It is a quantity which admits of being measured on a linear scale. For example, temperature is measured on the scale of a thermometer; angle on a graduated scale. A physical or geometric quantity is said to be a *scalar* if it is such that its instances (states) admit of a one-to-one reciprocal continuous correspondence with length. The real numbers are taken *per se* to be scalars. This assumption is equivalent to the coördinate axiom discussed in §1. A thermometer is an instrument which establishes a correspondence between states of temperature and numbers on a scale; a thermometer is, then, a coördinate system.

We take as our concept of a *vector* quantity a quantity whose instances (states) admit of a one-to-one reciprocal continuous correspondence with a set of translations in space. Hence a vector quantity requires three ordered numbers for its specification.

Let **V** be any given vector[7] and let *T* be the unique translation determined by it in the above correspondence. Let *A* be any point, and let *B* be its correspondent under the translation *T*. We have seen that the translation is unambiguously represented by an arrow having *A* as its initial point and *B* as its terminal point. Hence this same arrow likewise completely describes the vector **V**. Any two arrows (directed line segments) satisfying the relations (1), (2), (3) (see page 5) represent the same translation and hence the same vector. Thus any point may be taken as the initial point of an arrow which represents a given vector.

As an aid to our visualization process and intuition, this graphical representation of a vector is extremely useful.

[7] We follow the practice of denoting vectors by **boldface type**.

While it is true that most of the theorems which we will establish by means of vectors will be proved algebraically, the motivation of the proof will often be through the graphical representation of the vectors.

Exercises

2.1. Classify the following with regard to scalars, vectors, neither: linear displacement, mass, time, force, speed, affine transformation, linear velocity, density, resistance, linear acceleration, angular velocity, angular rotation, homogeneous strain.

2.2. Give examples of coördinate systems which you have observed.

§3. Addition and Scalar Multiplication of Vectors

3.1 Addition of vectors.

References: Addition and scalar multiplication of vectors will necessarily be treated in any book on the subject.

Let the translations T_1, T_2, T_3 of page 7 be represented by the vectors V_1, V_2, V_3, respectively. In vector notation we denote the following of one translation by another by the sign + (plus), the process being called *addition*. Thus the vector equivalent of

$$T_3 = T_2T_1 = T_1T_2$$

is

$$V_3 = V_1 + V_2 = V_2 + V_1.$$

Hence the addition of vectors is commutative.

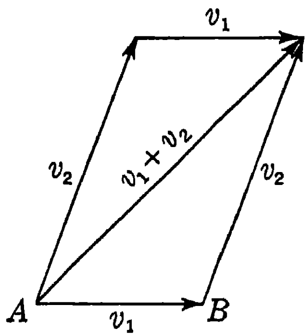

Fig. 6

We clearly have the graphical representation of the sum of two vectors: Let the arrow $\overrightarrow{AB}$, that is, the arrow from point A to point B, represent one vector, and let the arrow $\overrightarrow{BC}$ represent the other vector; then the sum of the vectors is represented by the arrow $\overrightarrow{AC}$ (Fig. 6).

If the arrows representing two vectors are not parallel we have the useful interpretation: The sum of two non-

parallel vectors is represented by a diagonal of a parallelogram whose sides represent the given vectors; the diagonal to be selected is the one which joins the initial point of one vector with the terminal point of the other vector (Fig. 6). This is the so-called *parallelogram law* for the sum or resultant of two vectors. The reader has probably encountered it in mechanics for compounding forces. However, in that connection the parallelogram law is not a theorem in the sense of mathematics, but only an empirical deduction (cf. *Mach* (43), pp. 33–48). So far as we are able to determine experimentally, forces behave as vectors. However, *if* forces *are* vectors, any theorem concerning vectors yields, through this interpretation, a statement concerning forces. It is through special interpretations such as this that the mathematics gains its utility with respect to other subjects. (Cf. *Veblen-Young* (51), I, pp. 1–2.)

Let T_1 be a translation and let T_2 be its *inverse*. Then the resultant of the two translations is the identity translation for the group which causes an arbitrary point to correspond to itself. Let $\boldsymbol{0}$, called the zero-vector, represent the identity translation. Then if V_1 represents a translation and V_2 its inverse, we have

$$V_1 + V_2 = \boldsymbol{0},$$

where we have emphasized in the notation that the right member is the zero-vector and not the number (scalar) zero. We see from the graphical representation that the arrows representing V_1 and V_2 differ only in sense of direction. We then write

$$V_2 = -V_1.$$

3.2 Scalar multiplication of a vector.

If T is a translation and λ is a positive integer, then by T^λ we shall mean that the translation T is carried out λ times. This would cause a displacement λ times as great as the displacement caused by T alone, the displacement

being in the same direction. More generally, if T denotes the translation

$$T:\quad \begin{aligned} x_1' &= x_1 + \alpha \\ x_2' &= x_2 + \beta \\ x_3' &= x_3 + \gamma, \end{aligned}$$

then by T^λ we shall understand the translation

$$T^\lambda:\quad \begin{aligned} x_1' &= x_1 + \lambda\alpha \\ x_2' &= x_2 + \lambda\beta \\ x_3' &= x_3 + \lambda\gamma \end{aligned}$$

where λ *is any real number*. We take this as a *definition* of λV. If V is the vector representing T, then we must take λV as representing T^λ; otherwise, multiplication of a vector by a positive integer would be inconsistent with addition.

Hence, if a vector V is represented by an arrow, the vector λV, where λ is any real number, will be represented by an arrow having the same or the opposite direction according as λ is positive or negative, and having the length $|\lambda|$ times the length of the arrow representing V.

3.3 Laws governing addition and scalar multiplication of vectors.

Reference: See, in particular, *Weyl* (53), pp. 15–16.

We now list the following laws as governing the operations of addition and scalar multiplication of vectors. Some of these may be obtained as theorems resulting from our previous assumptions; others are properly regarded as additional assumptions.

Let a, b, c, . . . be any vectors, and let λ, μ, . . . be any real numbers.

I. Addition.

(1) $a + b = b + a$ (Commutative law for addition)

(2) $(a + b) + c = a + (b + c)$ (Associative law for addition)

(3) If a and b are any two vectors, there exists one and only one vector x satisfying the equation

$$a + x = b \qquad \text{(Unique solvability)}$$

Clearly, these three properties may be regarded as theorems resulting from our previous considerations.

II. Multiplication.

(1) $\lambda a = a\lambda$ (Multiplication is commutative with respect to scalars)

(2) $(\lambda + \mu)a = \lambda a + \mu a$ (Scalar multiplication is distributive with respect to addition of scalars)

(3) $\lambda(\mu a) = (\lambda\mu)a$ (Associative with respect to scalar multiplication)

(4) $\lambda(a + b) = \lambda a + \lambda b$ (Distributive with respect to addition of vectors)

(5) $1a = a$ (Notation)

Of these we regard (1) as a new assumption and (5) as a convention; the remaining ones can be established as theorems. For the axioms underlying the real number system, see *Veblen-Young* (51), I, p. 149; *Dickson* (30), p. 200; *Beck* (25), Chapter I; *Carathéodory* (29), pp. 1–5.

III. Points and vectors.

(1) Every ordered pair of points A and B determine uniquely a vector a; expressed symbolically,

$$\overrightarrow{AB} = a.$$

(2) If A is any point and a is any vector, there exists a unique point B such that

$$\overrightarrow{AB} = a.$$

(3) If $\overrightarrow{AB} = a$ and $\overrightarrow{BC} = b$, then $\overrightarrow{AC} = a + b$.

Exercises

3.1. The set of real vectors constitutes a commutative group with respect to addition.

3.2. Prove property (3) under laws of addition stated above.

3.3. Prove properties (2), (3), (4) under laws of scalar multiplication stated above.

§4. Linear Dependence of Vectors

4.1 Linear spaces.

References: *Weyl* (53), p. 16; *Carathéodory* (29), Chapter VI; *Schreier-Sperner* (49), I, pp. 22–35; *Veblen* (50), p. 13.

If $\boldsymbol{a}_1$ is a vector different from $\boldsymbol{0}$, all vectors of the form $\lambda_1\boldsymbol{a}_1$, λ_1 being an arbitrary real number, are said to form a *one-dimensional linear vector space.* If $\boldsymbol{a}_2$ is a vector not of the form $\lambda_1\boldsymbol{a}_1$, then all vectors of the form

$$\lambda_1\boldsymbol{a}_1 + \lambda_2\boldsymbol{a}_2$$

constitute a *two-dimensional linear vector space.* If $\boldsymbol{a}_3$ is a vector not of the form $\lambda_1\boldsymbol{a}_1 + \lambda_2\boldsymbol{a}_2$, then all vectors given by

$$\lambda_1\boldsymbol{a}_1 + \lambda_2\boldsymbol{a}_2 + \lambda_3\boldsymbol{a}_3,$$

where the λ's are arbitrary real numbers, constitute a *three-dimensional linear vector space.*

If O is a fixed point and $\boldsymbol{a}_1$ is a vector $\neq \boldsymbol{0}$, the end points P of all vectors $\overrightarrow{OP}$ of the form $\lambda_1\boldsymbol{a}_1$, λ_1 being an arbitrary real number, constitute a *one-dimensional space of points* called a *straight line.* If $\boldsymbol{a}_2$ is a vector not of the form $\lambda_1\boldsymbol{a}_1$, the end points P of all vectors of the form $\overrightarrow{OP} = \lambda_1\boldsymbol{a}_1 + \lambda_2\boldsymbol{a}_2$ form a *two-dimensional space of points* called a *plane.* Thus the plane may be thought of as generated by sliding, in a particular fashion, one straight line along another. Similarly, the end points P of vectors $\overrightarrow{OP}$ of a three-dimensional linear vector space form a *three-dimensional space of points.*

4.2 Linear dependence.

References: *Carathéodory* (29), p. 309; *Schreier-Sperner* (49), I, pp. 19–22; *Weyl* (53), pp. 16–17; *Beck* (25), p. 74; *Bôcher* (27), Chapter III; *Graustein* (35), pp. 8–12.

A finite number of vectors $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_h$ are said to be *linearly dependent* with respect to the field of real numbers

λ if there exist real numbers $\lambda_1, \lambda_2, \ldots, \lambda_k$ not all zero, such that,

$$\lambda_1 a_1 + \lambda_2 a_2 + \cdots + \lambda_k a_k = 0$$

The vectors are said to be *linearly independent* if they are not linearly dependent. Hence, if a set of vectors known to be linearly independent satisfy an equation of the form

$$\lambda_1 a_1 + \lambda_2 a_2 + \cdots + \lambda_k a_k = 0,$$

it must be that $\lambda_1 = \lambda_2 = \cdots = \lambda_k = 0$. The notion of linear dependence as applied to vectors is indispensable if we wish to reach a clear understanding of vector analysis. Unfortunately, this essential idea has been frequently omitted or very inadequately treated by authors of books on the subject; however, it has been tacitly assumed in all such cases. Whenever we speak of linear dependence, it will always be understood to be with respect to the field of real numbers.

4.3 Axiom of dimensionality.

We now make the *assumption:* There exist n linearly independent vectors, but every set of $n + 1$ vectors are linearly dependent; unless it is otherwise indicated, we employ the further assumption that $n = 3$.

This means that we are restricting our attention to linear vector spaces of at most three dimensions. Also, all points P which are available belong to the same three-space of points.

4.4 Basis of a linear vector space.

References: *Weyl* (53), p. 17; *Schreier-Sperner* (49), I, p. 26; *Carathéodory* (29), p. 310.

We have postulated the existence of three linearly independent vectors. Let such a set be denoted by e_1, e_2, e_3. We have, also, assumed that any four vectors of the space under consideration are linearly dependent. Hence

for any vector V in the space, there exist real numbers λ, α_1, α_2, α_3 not all zero, such that

$$\lambda V + \alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3 = 0.$$

Here, $\lambda \neq 0$; otherwise there would exist a linear dependence among the vectors e_1, e_2, e_3 contrary to hypothesis. Hence we may write

$$V = \beta_1 e_1 + \beta_2 e_2 + \beta_3 e_3.$$

Hence, in the space considered, any vector of the form $\alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3$, where e_1, e_2, e_3 are linearly independent, is a vector of the space; and, conversely, any vector of the space can be so expressed. Any set of vectors having these two properties is called a *basis* of the space which they "span" or "generate." We shall make extensive use of a basis in our study of vector analysis. It turns out, usually, when we have a meaning for a given operation when it is applied to the base vectors, that then we know its meaning when it is applied to an arbitrary vector.

The linear independence or dependence of a set of vectors which are expressed as linear combinations of the base vectors can be characterized in terms of their coefficients. As an instance of this we prove the

Theorem: A necessary and sufficient condition that

$$a = \alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3 \text{ and } b = \beta_1 e_1 + \beta_2 e_2 + \beta_3 e_3$$

be linearly dependent, where e_1, e_2, e_3 constitute a basis, is that the rank of the matrix

$$\begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \end{pmatrix}$$

be less than 2; that is, that every second-order determinant formed from two of its columns be equal to zero.

By the definition of linear dependence, a necessary and sufficient condition that a and b be linearly dependent is that there exist numbers λ, μ not both zero such that

$$\lambda a + \mu b = 0.$$

In terms of the base vectors this condition becomes

$$\lambda(\alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3) + \mu(\beta_1 e_1 + \beta_2 e_2 + \beta_3 e_3) = 0$$

or

$$(\alpha_1\lambda + \beta_1\mu)e_1 + (\alpha_2\lambda + \beta_2\mu)e_2 + (\alpha_3\lambda + \beta_3\mu)e_3 = 0.$$

Since, by hypothesis, e_1, e_2, e_3 are linearly independent, the numbers λ, μ must satisfy the three linear algebraic (scalar) equations

$$(A) \qquad \begin{aligned} \alpha_1\lambda + \beta_1\mu &= 0 \\ \alpha_2\lambda + \beta_2\mu &= 0 \\ \alpha_3\lambda + \beta_3\mu &= 0. \end{aligned}$$

Of course, these equations are satisfied by the pair of numbers $\lambda = 0$, $\mu = 0$, no matter what the coefficients α, β are. However, our problem is to obtain a condition on the coefficients α, β which will be equivalent to the existence of a solution λ, μ, not both zero. If λ, μ are to satisfy this system of three equations, in particular, they must satisfy any two of the equations. Consider, then,

$$\begin{aligned} \alpha_1\lambda + \beta_1\mu &= 0 \\ \alpha_2\lambda + \beta_2\mu &= 0. \end{aligned}$$

If the determinant

$$\begin{vmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{vmatrix} \neq 0,$$

these equations can be solved by determinants in the usual manner and clearly yield the result $\lambda = 0$, $\mu = 0$. Hence we conclude that if the equations (A) have a solution other than 0, 0, the determinant of coefficients of any pair of these equations must vanish. Conversely, if these determinants are zero we verify that any solution of one of the equations is also a solution of each of the others.

4.5 Affine coördinate system.

References: *Schreier-Sperner* (49), I, p. 32; *Weyl* (53), p. 18.

Let $e_1, e_2, \ldots, e_n$, with n a positive integer, be n linearly independent vectors. Then the set of vectors for

which $e_1, e_2, \ldots, e_n$ form a basis is called a *linear vector space of n-dimensions*. If O is a fixed point, the set of end points P of vectors of the form

$$\overrightarrow{OP} = \alpha_1 e_1 + \alpha_2 e_2 + \cdots + \alpha_n e_n$$

is called an *n-dimensional space of points*. The point O together with the vectors $e_1, e_2, \ldots, e_n$ is called an *affine coördinate system* for the space of points. The numbers of the ordered set $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ corresponding to the point P, the terminal point of the vector $\overrightarrow{OP}$, are known as the *coördinates* of the point P. For $n = 1, 2, 3$, we have, respectively, an affine coördinate system for a line, a plane, and a three-space of points, as previously presented.

There is nothing mysterious concerning an n-dimensional space as used in the mathematical sense when properly understood. We are careful to avoid the term "space" with any meaning other than is implied by the above definitions. Thus, we are not here interested in the dimensionality of any physical, intuitive, psychological, or any kind of space other than mathematical. See *Weyl* (53), p. 23.

Exercises

4.1. A necessary and sufficient condition that two vectors be linearly dependent is that they be parallel; that is, that their graphical representations be parallel.

4.2. A necessary and sufficient condition that three vectors be linearly dependent is that they be parallel to a plane.

4.3. Assuming that e_1, e_2, e_3 constitute a basis, determine which of the following sets of vectors are linearly dependent:

(1) $a = 2e_1 - e_2 - 2e_3$
$b = 6e_1 - 3e_2 - 6e_3$

(2) $a = 2e_1 - e_2 - 2e_3$
$b = 6e_1 - 3e_2 + e_3$

(3) $a = 2e_1 - e_2 + 7e_3$
$b = e_1 + 4e_2 + 11e_3$
$c = 3e_1 - 6e_2 + 3e_3$

(4) $a = 3e_1 + e_2 + e_3$
$b = 2e_1 + 5e_2 - e_3$
$c = 6e_1 + 17e_2 - e_3.$

4.4. A necessary and sufficient condition that a vector be the zero-vector is that it have coefficients 0, 0, 0 when expressed in terms of a basis.

4.5. A necessary and sufficient condition that a, b, c be linearly dependent, when expressed in terms of a basis

$$\begin{aligned} a &= \alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3 \\ b &= \beta_1 e_1 + \beta_2 e_2 + \beta_3 e_3 \\ c &= \gamma_1 e_1 + \gamma_2 e_2 + \gamma_3 e_3, \end{aligned}$$

is that the rank of the matrix of coefficients be less than 3; that is, that the determinant

$$\begin{vmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \\ \gamma_1 & \gamma_2 & \gamma_3 \end{vmatrix}$$

shall vanish.

4.6. Determine α_1, α_2, α_3, such that the vectors

$$\begin{aligned} a &= \alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3 \\ b &= 2e_1 - e_2 + 7e_3 \\ c &= 3e_1 - 6e_2 + 3e_3 \end{aligned}$$

shall be linearly dependent. Characterize the situation geometrically.

4.7. A necessary and sufficient condition that three vectors of a three-dimensional linear vector space constitute a basis for the space is that they be linearly independent.

4.8. Let a be a vector in a space having e_1, e_2, e_3 as a basis. Show that the resolution

$$a = \alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3$$

is unique; that is, that the coefficients α_1, α_2, α_3 are unique.

4.9. The vectors of any linear vector space are closed with respect to addition and scalar multiplication; that is, if x and y belong to a linear vector space L, then $x + y$ and λx also belong to L.

4.10. Prove the following theorems:

(1) If a sub-set of p vectors is linearly dependent, then the p vectors are likewise linearly dependent.

(2) If p vectors are linearly independent, then any sub-set of those p vectors is also linearly independent.

(3) If p vectors are linearly dependent, $p > 1$, then at least one of the vectors is a linear combination of the remaining ones.

(4) If the vectors $a_1, a_2, \ldots, a_p$ are linearly independent but the vectors $a_1, a_2, \ldots, a_p, b$ are linearly dependent, then b is a linear combination of $a_1, a_2, \ldots, a_p$.

(5) If the vectors $a_1, a_2, \ldots, a_p$ are linearly independent and b is not a linear combination of them, then the vectors $a_1, a_2, \ldots, a_p, b$ are also linearly independent.

(6) Any set of vectors containing the zero-vector is a linearly dependent set.

§5. Illustrative Examples

We are now in position to work a variety of problems. It will be noted that the results are obtained by means of affine coördinates, and they are in each case independent of the particular coördinate system used. Hence the results hold if the coördinate system chosen is a rectangular Cartesian one; however, no advantage is gained by employing such a specialized coördinate system. The results are partially characterized by the entire absence of any metrical relations such as the *distance* between two points or the *angle* between two lines.

Example 1. Show that the points whose affine coördinates are $(5, -1)$, $(-1, 2)$, $(-5, 0)$, and $(1, -3)$ are the vertices of a parallelogram. It will be sufficient to show that $\overrightarrow{BA} = \overrightarrow{CD}$.

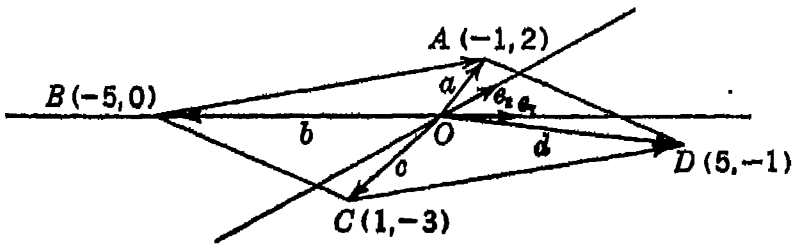

Fig. 7

Clearly $b + \overrightarrow{BA} = a$ and $c + \overrightarrow{CD} = d$, and hence $\overrightarrow{BA} = a - b$ and $\overrightarrow{CD} = d - c$. But

$$\begin{aligned} a &= -e_1 + 2e_2 \\ b &= -5e_1 \\ c &= e_1 - 3e_2 \\ d &= 5e_1 - e_2. \end{aligned}$$

Therefore, $\overrightarrow{BA} = 4e_1 + 2e_2$ and $\overrightarrow{CD} = 4e_1 + 2e_2$; that is,

$$\overrightarrow{BA} = \overrightarrow{CD}.$$

Hence we conclude that the figure is a parallelogram.

It is important to appreciate that the result is independent of the particular affine coördinate system used. In other words, if two independent workers constructed the figures, the figures would probably "look" quite different. In each case, however, the constructed figure could not fail to be a parallelogram, which is all that we are asked to prove.

Example 2. The diagonals of a parallelogram bisect one another.

Let vectors a, b be the adjacent sides of a parallelogram, with c and d its diagonals as indicated in Fig. 8.

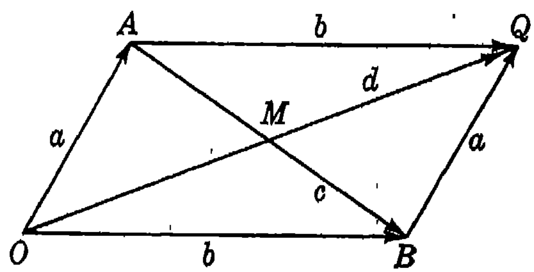

Fig. 8

Then $d = a + b$, and $a + c = b$ or $c = b - a$. Let M be the point of intersection of the diagonals, and consider the vector $\overrightarrow{OM}$. Since M is on the vector $\overrightarrow{OQ} = d$, it must be that

$$\overrightarrow{OM} = \lambda d,$$

where λ is a scalar yet to be determined. Similarly, since M is on the diagonal $\overrightarrow{AB} = \boldsymbol{c}$, we have $\overrightarrow{OM} = \boldsymbol{a} + \mu\boldsymbol{c}$ where μ is a suitable scalar. Equating these expressions for $\overrightarrow{OM}$, we have a condition on the unknown scalars λ, μ,

$$\lambda\boldsymbol{d} = \boldsymbol{a} + \mu\boldsymbol{c}.$$

This is a linear equation in the vectors $\boldsymbol{a}$, $\boldsymbol{c}$, and $\boldsymbol{d}$. Since, however, these vectors are *not linearly independent*, we can draw no immediate conclusion concerning values of λ and μ which satisfy this equation. We now express the vectors in this equation in terms of two vectors known to be linearly independent, say, $\boldsymbol{a}$ and $\boldsymbol{b}$. Thus we have the equation

$$\lambda(\boldsymbol{a} + \boldsymbol{b}) = \boldsymbol{a} + \mu(\boldsymbol{b} - \boldsymbol{a})$$

or

$$(\lambda + \mu - 1)\boldsymbol{a} + (\lambda - \mu)\boldsymbol{b} = \boldsymbol{0}.$$

Since $\boldsymbol{a}$ and $\boldsymbol{b}$ are linearly independent (otherwise they would not form the adjacent sides of a parallelogram), it must be that

$$\begin{aligned} \lambda + \mu - 1 &= 0 \\ \lambda - \mu &= 0. \end{aligned}$$

The solution of these equations for the unknown scalars λ and μ is

$$\lambda = \tfrac{1}{2},\ \mu = \tfrac{1}{2}$$

Hence $\overrightarrow{OM} = \frac{1}{2}\overrightarrow{OQ}$ and $\overrightarrow{AM} = \frac{1}{2}\overrightarrow{AB}$; that is, the point M bisects each of the diagonals.

Example 3. To divide a directed line segment $\overrightarrow{AB}$ into a prescribed ratio $\lambda : \mu$.

Denote the vector $\overrightarrow{AB}$ by $\boldsymbol{c}$; let O be a point not on the line AB; and let $\boldsymbol{a} = \overrightarrow{OA}$ and $\boldsymbol{b} = \overrightarrow{OB}$. Let P be the desired point, and let $\overrightarrow{OP} = \boldsymbol{p}$. Then, clearly,

$$\boldsymbol{p} = \boldsymbol{a} + \frac{\lambda}{\lambda + \mu}\boldsymbol{c} = \boldsymbol{a} + \frac{\lambda}{\lambda + \mu}(\boldsymbol{b} - \boldsymbol{a})$$

or

$$\boldsymbol{p} = \frac{\mu\boldsymbol{a} + \lambda\boldsymbol{b}}{\lambda + \mu},$$

the required solution Since the resolution of a vector in terms of a basis is unique, this single vector equation is equivalent to

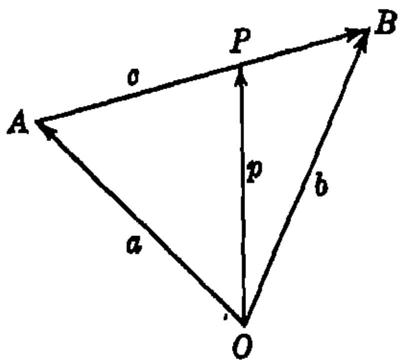

Fig. 9

three (or n) scalar equations Let O be the origin of an affine coördinate system determined by base vectors e_1, e_2, e_3. Let

$$
\begin{aligned}
a &= \alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3 \\
b &= \beta_1 e_1 + \beta_2 e_2 + \beta_3 e_3 \\
p &= \pi_1 e_1 + \pi_2 e_2 + \pi_3 e_3.
\end{aligned}
$$

Then the coordinates of A, B, P are, respectively, $(\alpha_1, \alpha_2, \alpha_3)$, $(\beta_1, \beta_2, \beta_3)$, (π_1, π_2, π_3). Hence the coördinates of P are given by

$$\pi_1 = \frac{\mu\alpha_1 + \lambda\beta_1}{\lambda + \mu}, \qquad \pi_2 = \frac{\mu\alpha_2 + \lambda\beta_2}{\lambda + \mu}, \qquad \pi_3 = \frac{\mu\alpha_3 + \lambda\beta_3}{\lambda + \mu},$$

the usual coördinate form

Example 4 Obtain the equation of the straight line passing through a point A and parallel to a given line.

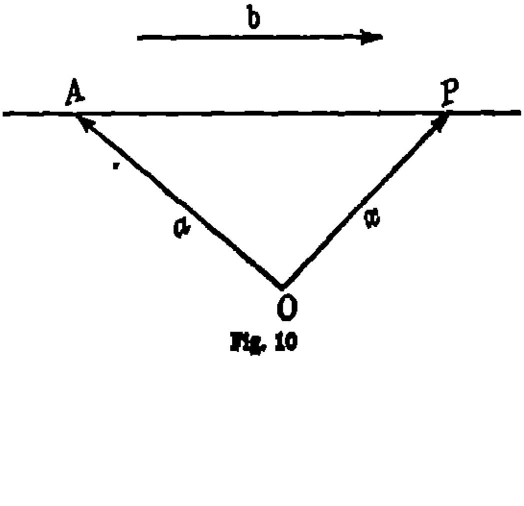

Fig. 10

Let O be any point, and denote $\overrightarrow{OA}$ by a; also let $b \neq 0$ be parallel to the given line. Let x be the vector $\overrightarrow{OP}$ where P is an arbitrary point on the required line. Then

$$x = a + \overrightarrow{AP} = a + \lambda b,$$

where λ is a variable scalar. We call this a vector equation of the required line, since any point P on the line determines a vector $x = \overrightarrow{OP}$ which satisfies this equation; and conversely, any x satisfying this equation determines a point P which is on the required line. Let O, e_1, e_2, e_3 be an affine coordinate system with origin at O in which

$$\begin{aligned} a &= \alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3 \\ b &= \beta_1 e_1 + \beta_2 e_2 + \beta_3 e_3 \\ x &= x_1 e_1 + x_2 e_2 + x_3 e_3. \end{aligned}$$

Then

$$x_1 e_1 + x_2 e_2 + x_3 e_3 = (\alpha_1 + \lambda\beta_1)e_1 + (\alpha_2 + \lambda\beta_2)e_2 + (\alpha_3 + \lambda\beta_3)e_3$$

Since the vectors e_1, e_2, e_3 are linearly independent, this vector equation is equivalent to the three scalar equations

$$\begin{aligned} x_1 &= \alpha_1 + \beta_1\lambda \\ x_2 &= \alpha_2 + \beta_2\lambda \\ x_3 &= \alpha_3 + \beta_3\lambda, \end{aligned}$$

which are the parametric equations of the line with λ the variable parameter. If λ is eliminated from these equations, we obtain the symmetric equations of the line

$$\frac{x_1 - \alpha_1}{\beta_1} = \frac{x_2 - \alpha_2}{\beta_2} = \frac{x_3 - \alpha_3}{\beta_3}.$$

Since by hypothesis $b \neq 0$, not all β_1, β_2, β_3 are zero.

Exercises

5.1. The lines joining the midpoints of the adjacent sides of any quadrilateral form a parallelogram.

5.2. Show by means of vectors that the medians of a triangle meet in a point which is a point of trisection of each median.

5.3. Show by means of vectors that the affine coordinates of the centroid of the triangle whose vertices are $(\alpha_1, \alpha_2, \alpha_3)$, $(\beta_1, \beta_2, \beta_3)$, $(\gamma_1, \gamma_2, \gamma_3)$ are

$$\left(\frac{\alpha_1 + \beta_1 + \gamma_1}{3}, \frac{\alpha_2 + \beta_2 + \gamma_2}{3}, \frac{\alpha_3 + \beta_3 + \gamma_3}{3}\right).$$

5.4. Obtain the vector equation of the straight line passing through two points A, B whose position vectors are $\boldsymbol{a}$ and $\boldsymbol{b}$, respectively. From the vector equation obtain the usual coordinate form. *Note* If A is a point and $\overrightarrow{OA} = \boldsymbol{a}$, the vector $\boldsymbol{a}$ is called the position vector of A (with respect to the point O).

5.5. Obtain a vector equation of a plane which passes through a point C whose position vector is $\boldsymbol{c}$ and which is parallel to two linearly independent vectors $\boldsymbol{a}$ and $\boldsymbol{b}$. Obtain the coordinate form.

5.6. If $\alpha\boldsymbol{a} + \beta\boldsymbol{b} = \boldsymbol{0}$ and $\alpha + \beta = 0$, with α, β not both zero, show that $\boldsymbol{a}$ and $\boldsymbol{b}$ are equal vectors.

5.7. If $\alpha\boldsymbol{a} + \beta\boldsymbol{b} + \gamma\boldsymbol{c} = \boldsymbol{0}$ and $\alpha + \beta + \gamma = 0$, with α, β, γ not all zero, show that if $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c}$ have a common point as their initial point, their terminal points lie on a straight line.

5.8. If ABC is a triangle, show that there exists a triangle whose sides are respectively equal to and parallel to the medians of the triangle ABC.

5.9. If a system of forces acts upon a particle, what is a condition upon the vectors representing the forces that the particle be in equilibrium?

5.10. Let O be any point in the plane of a triangle ABC. Let the midpoints of the sides of the triangle be A', B', C'. Then

$$\overrightarrow{OA} + \overrightarrow{OB} + \overrightarrow{OC} = \overrightarrow{OA'} + \overrightarrow{OB'} + \overrightarrow{OC'}$$

Various other applications will be found in the references cited. Here we are not interested in the results of the applications as such, but only insofar as they contribute toward an understanding of the subject of vector analysis and indicate its significance in related fields of knowledge.

§6. Introduction of a Metric

Thus far our study has been nonmetrical in character. We have had no occasion to speak of the *distance* between

two points, or of the *length* of an arbitrary vector, or of the *angle* between two vectors. It seems appropriate at this time to introduce a metric. By a *metric* is meant a scheme for measuring distances.

6.1 **Fundamental quadratic form.**

References: *Weyl* (53), pp. 22–33, *Cartan* (28), p. 29, *Böcher* (27), Chapter XI, *Kowalewski* (41), p. 176 *ff.*

Let e_1, e_2, e_3 constitute a basis for a linear vector space L_3 of three dimensions. A point O together with these vectors forms an affine coordinate system for the associated three-space of points E_3. In addition to our previous assumptions, let there now be given a *quadratic form*

$$\begin{aligned} Q(x) = \lambda_{11}x_1^2 + \lambda_{12}x_1x_2 + \lambda_{13}x_1x_3 \\ + \lambda_{21}x_2x_1 + \lambda_{22}x_2^2 + \lambda_{23}x_2x_3 \\ + \lambda_{31}x_3x_1 + \lambda_{32}x_3x_2 + \lambda_{33}x_3^2 \end{aligned}$$

with the properties

(1) $\lambda_{\alpha\beta}$ are real constants,
(2) $\lambda_{\alpha\beta} = \lambda_{\beta\alpha}$,
(3) $Q(x) > 0$ for all real numbers x_1, x_2, x_3, not all zero.

On account of (1) the quadratic form is said to be *real*; on account of (2) the form is said to be *symmetric*, because of (3) it is said to be *positive definite*.

Given the quadratic form $Q(x)$, which we now write in the more convenient notation

$$Q(x) = \sum_{\alpha,\beta=1}^{3} \lambda_{\alpha\beta}x_\alpha x_\beta,$$

we also have the uniquely determined *symmetric bilinear form*

$$Q(x, y) = \sum_{\alpha,\beta=1}^{3} \lambda_{\alpha\beta}x_\alpha y_\beta.$$

Now, given any two vectors x and y in L_3, $x = x_1e_1 + x_2e_2 + x_3e_3$ and $y = y_1e_1 + y_2e_2 + y_3e_3$, we *define the*

length of the vector x *to be the non-negative real number* $+\sqrt{Q(x)}$, *and the* angle θ between the vectors x and y, *neither of which is the zero-vector*, *by*

$$\cos\theta = \frac{Q(x, y)}{\sqrt{Q(x)}\sqrt{Q(y)}}.$$

The resulting geometry is called *Euclidean*.

Evidently

$$\begin{aligned} Q(x, y) &= Q(y, x) \\ Q(x + y, z) &= Q(x, z) + Q(y, z) \\ Q(\lambda x, y) &= \lambda Q(x, y), \end{aligned}$$

λ being an arbitrary real number.

We now prove the important result. If x and y are any two non-zero real vectors of L_3, the angle θ between them as above defined is real, that is, not imaginary. Obviously the expression for $\cos\theta$ is real. It will be sufficient, then, to show that the absolute value of

$$\frac{Q(x, y)}{\sqrt{Q(x)}\sqrt{Q(y)}}$$

does not exceed 1, or that

$$|Q(x, y)| \leq \sqrt{Q(x)}\sqrt{Q(y)},$$

or that

$$\{Q(x, y)\}^2 \leq Q(x)Q(y).$$

Consider, then, $Q(x + \lambda y)$ where λ is an arbitrary scalar. We have

$$Q(x + \lambda y) = Q(x) + 2Q(x, y)\lambda + Q(y)\lambda^2,$$

a quadratic expression in λ with real coefficients. By hypothesis $Q(x + \lambda y) \geq 0$, the equality sign holding only when $x + \lambda y = 0$. Hence the *equation* in λ

$$Q(y)\lambda^2 + 2Q(x, y)\lambda + Q(x) = 0$$

possesses no distinct real roots. But a necessary and sufficient condition that this be true is that the discriminant of the quadratic expression be negative or zero; that is, that

$$\{Q(x, y)\}^2 - Q(y)Q(x) \leq 0.$$

This completes the proof of the theorem

The important relation

$$[Q(x, y)]^2 \leq Q(x)Q(y)$$

is known as the *Cauchy-Schwarz inequality*.

If $Q(x) = 1$, the vector x is said to be unitary or to be a unit vector. Nothing more is to be implied by this terminology than that the vector has the length 1 (one) If x and y are each non-zero vectors, the vanishing of $Q(x, y)$ is equivalent to $\cos\theta = 0$. Hence, *a necessary and sufficient condition that two non-zero vectors x and y be perpendicular, or orthogonal, is that* $Q(x, y) = 0$. It is clear that, if x is the zero-vector, $Q(x, y) = 0$ In this sense it will be convenient to speak of the zero-vector as being orthogonal to y.

6.2 Scalar product of two vectors.

The value of $Q(x, y)$ is called the *scalar product* of the vectors x and y. We shall denote the scalar product by

$$x \cdot y,$$

a notation probably due to Gibbs. This symbol is frequently read "x dot y" and is often called the *dot product* It is of course not a "product" as in algebra We have simply in $x \cdot y$ an operation which when applied to two vectors yields a scalar. In this notation the length of a vector x is given by $+\sqrt{x \cdot x}$. We shall also use the symbol $|x|$ to stand for the length of x.

Exercises

6.1. A necessary and sufficient condition that a real vector x be the zero-vector is that $x \cdot x = 0$

6.2. If $x \neq 0$, there exists a unique real positive number λ such that λx is a unit vector

6.3. If x and y are unit vectors, the angle θ between them is given by $\cos\theta = x \cdot y$.

6.4. If a vector is orthogonal to each of three linearly independent vectors, it is the zero-vector.

6.5. The scalar product of two vectors admits of the interpretation: the product of the length of one of the vectors and the orthogonal projection of the other vector upon it. (This is frequently taken as a definition of the scalar product of two vectors.)

6.6. Establish the following:

(1) $x \cdot y = y \cdot x$

(2) $z \cdot (x + y) = z \cdot x + z \cdot y$

(3) $|-x| = |x|$

(4) $|\lambda x| = |\lambda||x|$.

6.7. Show that

$$(x + y) \cdot (x + y) = x \cdot x + 2x \cdot y + y \cdot y$$

What well-known formula of trigonometry is this?

6.8. If x represents a displacement and y represents a force, then $x \cdot y$ gives the work done in displacing y through the displacement x. With this interpretation, what is the significance of property (2) of Exercise 6.6?

6.9. A necessary and sufficient condition that a parallelogram shall be a rhombus is that its diagonals intersect at right angles.

6.10. The distance between two points is taken to be the length of the vector joining them. Deduce the formula for the distance between two points in terms of their coordinates.

6.11. For any three points A, B, C of E_3,

$$\overline{AB} + \overline{BC} \geq \overline{AC},$$

where $\overline{AB}$ denotes the distance between A and B. This relation is frequently called the "triangle property."

6.12. Given the fundamental quadratic form $Q(x)$ as follows: $\lambda_{11} = 14$, $\lambda_{12} = -2$, $\lambda_{13} = 7$, $\lambda_{22} = 3$, $\lambda_{23} = 0$, $\lambda_{33} = 6$, compute the lengths of

$$x = 2e_1 - 2e_2 + e_3 \text{ and } y = e_1 + 4e_2 + e_3$$

and the angle between them. If x and y are the position vectors of A and B, respectively, compute the distance $\overline{AB}$. Verify the Law of Cosines for the triangle OAB.

6.13. Given base vectors e_1, e_2, e_3 for L_3 and the fundamental quadratic form

$$Q(x) = \sum_{\alpha,\beta=1}^{3} \lambda_{\alpha\beta} x_\alpha x_\beta,$$

establish the following

(1) $e_\alpha \cdot e_\beta = \lambda_{\alpha\beta}$

(2) $\lambda_{11}, \lambda_{22}, \lambda_{33}$ are each positive

(3) The cofactors of $\lambda_{11}, \lambda_{22}, \lambda_{33}$ in the determinant

$$g = \begin{vmatrix} \lambda_{11} & \lambda_{12} & \lambda_{13} \\ \lambda_{21} & \lambda_{22} & \lambda_{23} \\ \lambda_{31} & \lambda_{32} & \lambda_{33} \end{vmatrix}$$

are each positive

(4) The determinant g is positive (For this part of the exercise reference may be made to the following section.)

6.3 The i, j, k system of base vectors

References: *Cartan* (28), p. 30, *Bôcher* (27), p. 131, *Kowalewski* (41), p. 179.

We now show that there exists in L_3 a system of base vectors $\bar{e}_1, \bar{e}_2, \bar{e}_3$ which are unitary and mutually orthogonal. Let the given basis for L_3 be e_1, e_2, e_3 and let the quadratic form be

$$Q(x) = \sum_{\alpha,\beta=1}^{3} \lambda_{\alpha\beta} x_\alpha x_\beta$$

Since e_1, e_2, e_3 are linearly independent, no one of them is the zero-vector. Hence, in particular, $e_1 \cdot e_1 = \lambda_{11} \neq 0$. We first set up an orthogonal system of base vectors e_1', e_2', e_3'. Let

$$e_1' = e_1$$
$$e_2' = \alpha e_1 + \beta e_2.$$

Upon imposing the condition of orthogonality, $e_1' \cdot e_2' = 0$, α and β are required to satisfy the equation

$$\alpha e_1 \cdot e_1 + \beta e_1 \cdot e_2 = 0$$

or

$$\frac{\alpha}{\beta} = -\frac{e_1 \cdot e_2}{e_1 \cdot e_1}.$$

Hence the vector

$$e_2' = (-e_1 \cdot e_2)e_1 + (e_1 \cdot e_1)e_2 = \begin{vmatrix} e_1 \cdot e_1 & e_1 \\ e_2 \cdot e_1 & e_2 \end{vmatrix}$$

is perpendicular to e_1'. Now define e_3' by

$$e_3' = \begin{vmatrix} e_1 \cdot e_1 & e_1 \cdot e_2 & e_1 \\ e_2 \cdot e_1 & e_2 \cdot e_2 & e_2 \\ e_3 \cdot e_1 & e_3 \cdot e_2 & e_3 \end{vmatrix}.$$

Clearly $e_3' \cdot e_1' = 0$ and $e_3' \cdot e_2' = 0$. Hence e_1', e_2', e_3' form an orthogonal system. To obtain a unitary system, it is necessary only to divide each vector by its length. Hence the desired vectors $\bar{e}_1$, $\bar{e}_2$, $\bar{e}_3$ are given by

$$\bar{e}_1 = \frac{e_1'}{\sqrt{e_1 \cdot e_1}}$$

$$\bar{e}_2 = \frac{e_2'}{\sqrt{\begin{vmatrix} e_1 \cdot e_1 & e_1 \cdot e_2 \\ e_2 \cdot e_1 & e_2 \cdot e_2 \end{vmatrix}(e_1 \cdot e_1)}}$$

$$\bar{e}_3 = \frac{e_3'}{\sqrt{\begin{vmatrix} e_1 \cdot e_1 & e_1 \cdot e_2 & e_1 \cdot e_3 \\ e_2 \cdot e_1 & e_2 \cdot e_2 & e_2 \cdot e_3 \\ e_3 \cdot e_1 & e_3 \cdot e_2 & e_3 \cdot e_3 \end{vmatrix} \begin{vmatrix} e_1 \cdot e_1 & e_1 \cdot e_2 \\ e_2 \cdot e_1 & e_2 \cdot e_2 \end{vmatrix}}}.$$

We denote vectors of such a system by $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$, a well-standardized notation. A point O together with an $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ system of vectors constitutes a *rectangular Cartesian* coördinate system. The coefficients of the fundamental quadratic form for an $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ system are

$$\lambda_{11} = \lambda_{22} = \lambda_{33} = 1, \qquad \lambda_{12} = \lambda_{23} = \lambda_{31} = 0.$$

Exercises

6.14. If $x = x_1\boldsymbol{i} + x_2\boldsymbol{j} + x_3\boldsymbol{k}$ and $y = y_1\boldsymbol{i} + y_2\boldsymbol{j} + y_3\boldsymbol{k}$, then

$$|x| = \sqrt{x_1^2 + x_2^2 + x_3^2}$$

and

$$x \cdot y = x_1y_1 + x_2y_2 + x_3y_3.$$

6.15. Introduce an i, j, k system of base vectors, given the metric of Exercise **6.12**, and express the vectors x and y of that exercise in terms of the new base vectors. Verify the results of Exercise **6.12** by recomputing the same quantities in terms of the new metric.

6.16. Any mutually orthogonal, non-zero vectors are necessarily linearly independent.

6.17. Obtain the unique one-dimensional linear vector space L_1 which is orthogonal to the two-dimensional linear vector space L_2 generated by

$$x = x_1 e_1 + x_2 e_2 + x_3 e_3 \text{ and } y = y_1 e_1 + y_2 e_2 + y_3 e_3.$$

§7. Linear Transformations

References: *Weyl* (53), pp. 21–23; *Bôcher* (27), Chapter VI; *Schreier-Sperner* (49), I, pp. 114–128.

Let e_1, e_2, e_3 constitute a basis for L_3 having the fundamental quadratic form $Q(x) = \Sigma\lambda_{\alpha\beta}x_\alpha x_\beta$. We then have a meaning for the length of any vector in L_3 and the angle between any two such vectors. If now we introduce a new system of base vectors for the same space, let us investigate what modification of the quadratic form will insure that the length of an arbitrarily given vector, and the angle between two such vectors, shall be independent of the particular base system used.

7.1 Affine transformations.

Let new base vectors $\bar{e}_1$, $\bar{e}_2$, $\bar{e}_3$ be introduced by means of the equations

$$(A) \qquad \begin{aligned} \bar{e}_1 &= \alpha_{11}e_1 + \alpha_{12}e_2 + \alpha_{13}e_3 \\ \bar{e}_2 &= \alpha_{21}e_1 + \alpha_{22}e_2 + \alpha_{23}e_3 \\ \bar{e}_3 &= \alpha_{31}e_1 + \alpha_{32}e_2 + \alpha_{33}e_3, \end{aligned}$$

where the coefficients α_{ij} are constants with the determinant

$$\Delta \equiv \begin{vmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{vmatrix} \neq 0,$$

since the vectors are to be linearly independent. Let the

coefficients of the new quadratic form be denoted by $\bar{\lambda}_{\alpha\beta}$. Then, since $\bar{\lambda}_{\mu\nu} = \bar{e}_\mu \cdot \bar{e}_\nu$, we have

$$(B) \qquad \bar{\lambda}_{\mu\nu} = \sum_{\rho,\sigma=1}^{3} \lambda_{\rho\sigma}\alpha_{\mu\rho}\alpha_{\nu\sigma}.$$

Let x and y be any two vectors of L_3.

$$\begin{aligned} x &= x_1e_1 + x_2e_2 + x_3e_3 = \bar{x}_1\bar{e}_1 + \bar{x}_2\bar{e}_2 + \bar{x}_3\bar{e}_3 \\ y &= y_1e_1 + y_2e_2 + y_3e_3 = \bar{y}_1\bar{e}_1 + \bar{y}_2\bar{e}_2 + \bar{y}_3\bar{e}_3. \end{aligned}$$

The relations (A) between the base vectors imply the following relations between the coefficients of the vectors:

$$(C) \qquad \begin{aligned} x_1 &= \alpha_{11}\bar{x}_1 + \alpha_{21}\bar{x}_2 + \alpha_{31}\bar{x}_3 \\ x_2 &= \alpha_{12}\bar{x}_1 + \alpha_{22}\bar{x}_2 + \alpha_{32}\bar{x}_3 \\ x_3 &= \alpha_{13}\bar{x}_1 + \alpha_{23}\bar{x}_2 + \alpha_{33}\bar{x}_3. \end{aligned}$$

These same equations hold of course, for the y's and $\bar{y}$'s. Consider now

$$\begin{aligned} x \cdot y &= \sum_{\mu,\nu} \bar{\lambda}_{\mu\nu}\bar{x}_\mu\bar{y}_\nu \\ &= \sum_{\mu,\nu}\left\{\sum_{\rho,\sigma}\lambda_{\rho\sigma}\alpha_{\mu\rho}\alpha_{\nu\sigma}\bar{x}_\mu\bar{y}_\nu\right\} \\ &= \sum_{\rho,\sigma}\left\{\lambda_{\rho\sigma}\sum_{\mu,\nu}(\alpha_{\mu\rho}\bar{x}_\mu)(\alpha_{\nu\sigma}\bar{y}_\nu)\right\} \\ &= \sum_{\rho,\sigma}\lambda_{\rho\sigma}x_\rho y_\sigma. \end{aligned}$$

That is, *the scalar product $x \cdot y$ has the same value when computed in either system.* This exceedingly important result merits a formal statement. If two systems of base vectors are related by equations (A), the coefficients of an arbitrary vector x are related by equations (C). This is equivalent to what we mean by two descriptions

$$x_1e_1 + x_2e_2 + x_3e_3 \text{ and } \bar{x}_1\bar{e}_1 + \bar{x}_2\bar{e}_2 + \bar{x}_3\bar{e}_3$$

of the *same* vector. The two metrics are related by (B). The result we have established, then, is that the scalar

product of two vectors is *invariant* with respect to the simultaneous transformations (A), (B), and (C) on the base vectors, the metric, and the coefficients of a vector, respectively. Since the scalar product is invariant with respect to these transformations, it follows that the length of a vector, and the angle between two vectors, have values independent of the particular coördinate system used in the class of affine coördinate systems.

The importance of this result cannot be overemphasized. It means that a vector has an identity and significance independent of the coördinate system used. If it were not for this property, vectors could have no utility in physics or mechanics, as these subjects are primarily concerned with properties which are independent of the observer or of the particular apparatus (coördinate system) used in deducing the results. Suppose two observers M and N observe a vector quantity which they suspect is the *same* vector quantity, such as the flight of a meteor. Each observer has, of course, a coördinate system of his own. Let M's description of the vector be

$$x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + x_3\boldsymbol{e}_3$$

and that of N be $\bar{x}_1\bar{\boldsymbol{e}}_1 + \bar{x}_2\bar{\boldsymbol{e}}_2 + \bar{x}_3\bar{\boldsymbol{e}}_3$. The question as to whether or not they have observed the same vector quantity can have no meaning unless the relation between the two coördinate systems can be determined. Suppose the base vectors in the two systems are found to satisfy the relations (A); this means that the coefficients α_{11}, α_{12}, α_{13}, α_{23}, etc., are known. If then, the observed quantities x_1, x_2, x_3, and $\bar{x}_1$, $\bar{x}_2$, $\bar{x}_3$ satisfy the equations (C) with the known values of the coefficients α_{ij}, we say that they have observed the same vector quantity; otherwise not.

Equations of the form (A) or (C), which are linear equations with constant coefficients, are known as a *homogeneous affine transformation.* We note that if (A) is given, (B) and (C) are uniquely determined; also if (C) is given, then (A) and (B) are completely determined. If O is the common origin

of two affine coördinate systems determined by $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$, and $\bar{\boldsymbol{e}}_1$, $\bar{\boldsymbol{e}}_2$, $\bar{\boldsymbol{e}}_3$, the point P specified by a vector $\boldsymbol{x} = \overrightarrow{OP}$ has coördinates (x_1, x_2, x_3) in the one system and $(\bar{x}_1, \bar{x}_2, \bar{x}_3)$ in the other. Hence equations (C) may be interpreted as a transformation of one affine coördinate system to another having a common origin. That is, if the coördinates of a point P are given in one coördinate system, the equations (C) tell what the coördinates of the *same* point are in the other system.

7.2 Congruent transformation.

If the equations (A) are such that the quadratic form is transformed into itself; that is, such that $\bar{\lambda}_{\mu\nu} = \lambda_{\mu\nu}$, the transformation (A) is said to be a *congruent transformation.* In this case, since

$$\bar{\boldsymbol{e}}_\mu \cdot \bar{\boldsymbol{e}}_\nu = \bar{\lambda}_{\mu\nu} = \lambda_{\mu\nu} = \boldsymbol{e}_\mu \cdot \bar{\boldsymbol{e}}_\nu,$$

the new base vectors have respectively the same lengths as the original ones with the same angles between them. If the two sets $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$ and $\bar{\boldsymbol{e}}_1$, $\bar{\boldsymbol{e}}_2$, $\bar{\boldsymbol{e}}_3$ have the same orientation (discussed in §8), it means that one system of base vectors could be rotated into the other.

However, the equations (C) may be equally well interpreted as establishing a one-to-one reciprocal correspondence between the points of E_3 with themselves. To a point P having coördinates (x_1, x_2, x_3) there corresponds a unique point $\bar{P}$ with coördinates $(\bar{x}_1, \bar{x}_2, \bar{x}_3)$ in the *same coördinate system.* In the case of a congruent transformation, it results that the correspondents of P, Q, say $\bar{P}$, $\bar{Q}$, are such that the segment $\overline{PQ}$ has the same length as the segment $\overline{\bar{P}\bar{Q}}$.

7.3 Orthogonal transformation.

If a congruent transformation is carried out on an $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ system of vectors, it of course results in another $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ system. In this case the transformation is called an

orthogonal transformation. The equations for changing from one rectangular Cartesian coördinate system to another one with the same point as origin must then be an example of an orthogonal transformation.

The method used in **§6.1** of introducing a metric may seem strange on a first reading. A more familiar, but logically equivalent, approach to measurement of lengths of vectors and angles between them is achieved by the *postulate:* There exists a preferred coördinate system, called rectangular Cartesian, determined by a point O and three unit vectors $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$, mutually perpendicular to one another. In this system if

$$\boldsymbol{x} = x_1\boldsymbol{i} + x_2\boldsymbol{j} + x_3\boldsymbol{k} \text{ and } \boldsymbol{y} = y_1\boldsymbol{i} + y_2\boldsymbol{j} + y_3\boldsymbol{k},$$

the length of $\boldsymbol{x}$ is given by $\sqrt{x_1^2 + x_2^2 + x_3^2}$ and the angle θ between $\boldsymbol{x}$ and $\boldsymbol{y}$ is given by

$$\cos\theta = \frac{x_1y_1 + x_2y_2 + x_3y_3}{\sqrt{x_1^2 + x_2^2 + x_3^2}\sqrt{y_1^2 + y_2^2 + y_3^2}}.$$

Even though the wording somewhat disguises it, what we have actually supposed in this case is the particular quadratic form whose coefficients $\lambda_{\alpha\beta}$ are

$$\lambda_{11} = \lambda_{22} = \lambda_{33} = 1,\ \lambda_{12} = \lambda_{23} = \lambda_{31} = \lambda_{21} = \lambda_{32} = \lambda_{13} = 0,$$

by means of which we agree to measure lengths and angles in accordance with the above definitions. Starting with such a preferred coördinate system, if we introduce a new coördinate system by an orthogonal transformation, the quadratic form will remain unchanged. However, if we transform over to a new system by any affine transformation (A), the new quadratic form may be expected to have different coefficients; all we can definitely say about it is that it will have the properties (1), (2), (3) of **§6.1.**

Exercises

7.1. Starting with an $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ system of base vectors, let a new system of base vectors $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$ be introduced by equations (A).

Show that the new fundamental quadratic form will have the properties (1), (2), (3) of **§6.1.**

7.2. Starting with an $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ system, let new base vectors be given by

$$\begin{aligned} \boldsymbol{e}_1 &= \boldsymbol{i} - \boldsymbol{j} + 2\boldsymbol{k} \\ \boldsymbol{e}_2 &= \boldsymbol{i} + 2\boldsymbol{j} - \boldsymbol{k} \\ \boldsymbol{e}_3 &= 3\boldsymbol{i} + \boldsymbol{j} + \boldsymbol{k}. \end{aligned}$$

Compute the fundamental quadratic form for the new system. Given

$$\boldsymbol{x} = \boldsymbol{e}_1 - 2\boldsymbol{e}_2 + \boldsymbol{e}_3 \text{ and } \boldsymbol{y} = \boldsymbol{e}_1 + \boldsymbol{e}_2 + 2\boldsymbol{e}_3,$$

compute lengths of $\boldsymbol{x}$ and $\boldsymbol{y}$ and the angle between them. Express these vectors in terms of $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ and recompute the same quantities.

7.3. The set of all real affine transformations constitutes a group with respect to forming their resultant.

7.4. Let the equations

$$\begin{aligned} \bar{\boldsymbol{i}} &= \alpha_{11}\boldsymbol{i} + \alpha_{12}\boldsymbol{j} + \alpha_{13}\boldsymbol{k} \\ \bar{\boldsymbol{j}} &= \alpha_{21}\boldsymbol{i} + \alpha_{22}\boldsymbol{j} + \alpha_{23}\boldsymbol{k} \\ \bar{\boldsymbol{k}} &= \alpha_{31}\boldsymbol{i} + \alpha_{32}\boldsymbol{j} + \alpha_{33}\boldsymbol{k} \end{aligned}$$

define an *orthogonal* transformation. Prove the following:

(1) $\alpha_{\mu 1}^2 + \alpha_{\mu 2}^2 + \alpha_{\mu 3}^2 = 1$, $\mu = 1, 2, 3$.

(2) $\alpha_{\mu 1}\alpha_{\nu 1} + \alpha_{\mu 2}\alpha_{\nu 2} + \alpha_{\mu 3}\alpha_{\nu 3} = 0$, $\mu \neq \nu$.

(3) If Δ denotes the determinant of the tranformation, $\Delta^2 = 1$ and hence $\Delta = 1$ or -1.

(4) The coefficients $\alpha_{11}, \alpha_{12}, \alpha_{13}$ are the direction cosines of the vector $\bar{\boldsymbol{i}}$ with respect to the $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ system.

(5) The inverse transformation is

$$\begin{aligned} \boldsymbol{i} &= \alpha_{11}\bar{\boldsymbol{i}} + \alpha_{21}\bar{\boldsymbol{j}} + \alpha_{31}\bar{\boldsymbol{k}} \\ \boldsymbol{j} &= \alpha_{12}\bar{\boldsymbol{i}} + \alpha_{22}\bar{\boldsymbol{j}} + \alpha_{32}\bar{\boldsymbol{k}} \\ \boldsymbol{k} &= \alpha_{13}\bar{\boldsymbol{i}} + \alpha_{23}\bar{\boldsymbol{j}} + \alpha_{33}\bar{\boldsymbol{k}}. \end{aligned}$$

(6) If $\Delta = +1$, the transformation can be interpreted as representing a rigid displacement which leaves the origin fixed, that is, as a "rotation of axes."

(For orthogonal transformations, see *Schreier-Sperner* (49), I, p. 162.)

§8. Plane Areas as Vectors

References: *Gibbs-Wilson* (7), pp. 46–51; *Runge* (17), pp. 10–25.

In a Euclidean three-space of points E_3, there exists a unique direction, neglecting sense, which is perpendicular to any given plane. We may then conveniently represent a *plane area* by a vector.[8] Let R be a simply connected plane area bounded by a closed curve C. A region R is said to be *connected* if every two points in R admit of being joined by a continuous curve lying wholly in R; it is said to be *simply connected* if any two such curves joining the same two points admit of a continuous deformation which causes them to coincide. We assign to the bounding curve C a positive sense of direction (orientation), which we take as the positive sense of a tangent vector $\boldsymbol{t}$ to C at any point P on the curve (Fig. 11). Let $\boldsymbol{n}$ be a vector in the plane of R at P which is perpendicular to $\boldsymbol{t}$ and which joins P to a near-by point of R. The vector $\boldsymbol{n}$ is called an *inward-pointing normal.*

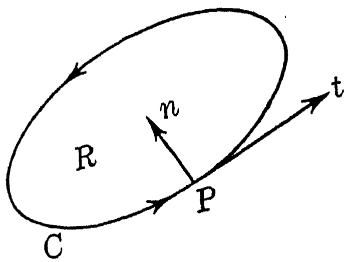

Fig. 11

We now take as a vector $\boldsymbol{r}$, to represent the plane area R, the vector which is uniquely determined by the three conditions:

(1) The vector is perpendicular to the plane of the area;
(2) Its length is equal to the value of the area;
(3) Its sense of direction is such that it forms with the tangent and normal vectors, $\boldsymbol{t}$ and $\boldsymbol{n}$, a right-handed triple in the order $\boldsymbol{t}$, $\boldsymbol{n}$, $\boldsymbol{r}$.

The positive sense of $\boldsymbol{r}$ is then the direction in which a right-handed screw would progress under a rotation from

[8] For a complete justification of this representation see *Runge* (17).

t toward n. Or, if the thumb of the right hand represents t and the index finger represents n, then the middle finger will represent r. The name "right-handed triple" comes from this description. If a man were walking along the shore of a lake with the water on his left, the vector representing the area would be pointing directly upward. If, however, the water were on his right, the vector representing it would be pointing directly downward.

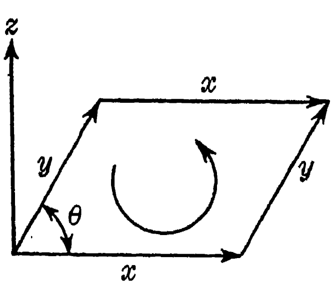

Fig. 12

8.1 Vector cross product of two vectors.

References: *Gibbs-Wilson* (7), p. 64 *ff.*; *Runge* (17), p. 34 *ff*.

Let x and y be the adjacent sides of a parallelogram, and let z be a vector representing the plane area of the parallelogram, which is oriented so that x, y, and z form a right-handed triple. Clearly, then, z is a function of x and y, say $z = \varphi(x, y)$, with the following properties:

(1) $\varphi(x, y) = -\varphi(y, x)$;
(2) $\varphi(x, x) = 0$;
(3) $\varphi(\lambda x, y) = \lambda\varphi(x, y)$;
(4) $\varphi(x, \lambda y) = \lambda\varphi(x, y)$.

Let θ be the angle between x and y. Then $z = \varphi(x, y)$ is such that

(5) $\sqrt{z \cdot z} = \sqrt{x \cdot x}\sqrt{y \cdot y} \sin\theta$.

In place of the notation $\varphi(x, y)$ we now use the well-standardized one

$$z = x \times y,$$

read "x cross y." The vector z is called the *vector cross product of x and y* in the order named.

The fundamental theorem

$$(x + y) \times z = (x \times z) + (y \times z)$$

may be established in a number of ways. We consider a proof which depends upon the notion of projected areas.

If a positive sense of the normal to a plane has been chosen, the plane is said to be *oriented.* Let C be a closed oriented curve bounding a simply connected region R in an oriented plane. The area of the region R is said to be positive or negative (relatively to the oriented plane) according as the vector representation of the plane area R has the same or opposite direction as the normal to the plane.

Let A' be the orthogonal projection on an oriented plane π' of a plane area A in an oriented plane π. The angle θ

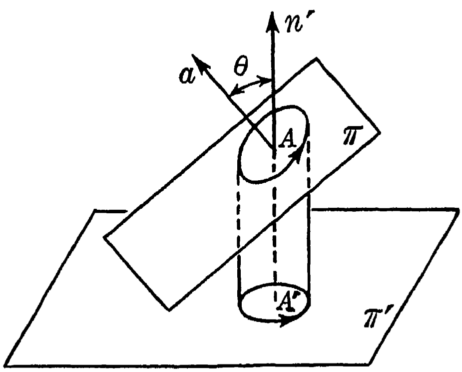

Fig. 13

between two oriented planes is defined as the angle between their (positive) normals. We have

$$(\text{area } A') = (\text{area } A) \cos \theta.$$

If $\boldsymbol{n}'$ is the unit normal vector of π' and $\boldsymbol{a}$ is the vector representation of the plane area A, the above can be expressed by

$$(\text{area } A') = \boldsymbol{a} \cdot \boldsymbol{n}'.$$

Consider now a region of space bounded by a closed surface S. Let $d\sigma$ be an element of surface area of S. We denote the vector representation of this surface element by $d\boldsymbol{\sigma}$. This is then a vector whose length is equal to $d\sigma$, and

which we take as having the direction of the outward-pointing normal to the surface S. The integral over S,

$$\int_S d\boldsymbol{\sigma}$$

is a vector, since it is the limit of a sum of vectors.

Let $\boldsymbol{a}$ be an arbitrary but fixed unit vector, and consider

$$\int_S d\boldsymbol{\sigma} \cdot \boldsymbol{a} = \boldsymbol{a} \cdot \int_S d\boldsymbol{\sigma}.$$

For simplicity we suppose that any line parallel to $\boldsymbol{a}$ which intersects the surface S will have exactly two points in

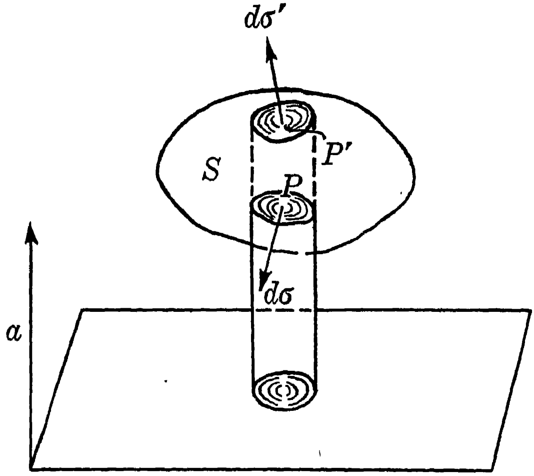

Fig. 14

common with S, or one point of tangency. Then to each point P of S there corresponds a unique point P' of S on the line through P parallel to $\boldsymbol{a}$. The projection of $d\sigma$ at P on a plane whose oriented normal is $\boldsymbol{a}$, is simply the numerical value of

$$\boldsymbol{a} \cdot d\boldsymbol{\sigma} \text{ or of } \boldsymbol{a} \cdot d\boldsymbol{\sigma}',$$

where the meaning of $d\sigma'$ is clear from the figure. But these are of opposite signs and hence their sum is zero. Thus we have

$$\boldsymbol{a} \cdot \int_S d\boldsymbol{\sigma} = 0,$$

and since $\mathbf{a}$ is arbitrary, we conclude that

$$\int_S d\mathbf{\sigma} = \mathbf{0}.$$

Now, given three linearly independent vectors $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$, let S be the closed surface formed by the faces of a prism, Fig. 15.

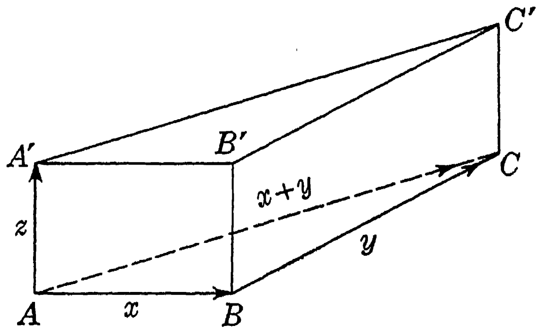

Fig. 15

Taking the outward-pointing normal, we have

$$\mathbf{x} \times \mathbf{z} = \int_{ABB'A'} d\mathbf{\sigma}$$

$$\mathbf{y} \times \mathbf{z} = \int_{BCC'B'} d\mathbf{\sigma}.$$

$$(\mathbf{x} + \mathbf{y}) \times \mathbf{z} = -\int_{ACC'A'} d\mathbf{\sigma}.$$

Clearly

$$\int_{ABC} d\mathbf{\sigma} + \int_{A'B'C'} d\mathbf{\sigma} = \mathbf{0}.$$

But by the above theorem

$$\int_S d\mathbf{\sigma} = \mathbf{0}$$

and hence we have

$$(\mathbf{x} \times \mathbf{z}) + (\mathbf{y} \times \mathbf{z}) - (\mathbf{x} + \mathbf{y}) \times \mathbf{z} = \mathbf{0};$$

that is,

$$(\mathbf{x} + \mathbf{y}) \times \mathbf{z} = (\mathbf{x} \times \mathbf{z}) + (\mathbf{y} \times \mathbf{z}),$$

which is the theorem we wished to establish. The reader will have no difficulty in proving this result for the case of coplanar vectors.

8.2 Linear operators.

Reference: *Burali-Forti and Marcolongo* (4), p. 16 *ff*.

Let L be an operator which operates on pairs of quantities to yield a quantity. It is said to be a *linear operator* if the following four properties hold:

$$
\begin{aligned}
&(1)\ L(x+z,\ y) = L(x,\ y) + L(z,\ y);\\
&(2)\ L(x,\ y+z) = L(x,\ y) + L(x,\ z);\\
&(3)\ L(\lambda x,\ y) = \lambda L(x,\ y),\ \lambda \text{ an arbitrary constant};\\
&(4)\ L(x,\ \lambda y) = \lambda L(x,\ y),\ \lambda \text{ an arbitrary constant}.
\end{aligned}
$$

Suppose $\boldsymbol{x}$ and $\boldsymbol{y}$ are the vectors

$$\boldsymbol{x} = x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + x_3\boldsymbol{e}_3;\ \boldsymbol{y} = y_1\boldsymbol{e}_1 + y_2\boldsymbol{e}_2 + y_3\boldsymbol{e}_3.$$

Then

$$
\begin{aligned}
L(\boldsymbol{x},\ \boldsymbol{y}) &= L(x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + x_3\boldsymbol{e}_3,\ y_1\boldsymbol{e}_1 + y_2\boldsymbol{e}_2 + y_3\boldsymbol{e}_3)\\
&= x_1y_1L(\boldsymbol{e}_1,\ \boldsymbol{e}_1) + x_1y_2L(\boldsymbol{e}_1,\ \boldsymbol{e}_2) + x_2y_1L(\boldsymbol{e}_2,\ \boldsymbol{e}_1)\\
&\quad + \cdots + x_3y_3L(\boldsymbol{e}_3,\ \boldsymbol{e}_3).
\end{aligned}
$$

Hence we shall have the meaning of any linear operator L operating on pairs of vectors $\boldsymbol{x}$, $\boldsymbol{y}$ when we know its meaning when operating on the various pairs of base vectors.

Consider now $\boldsymbol{x} \times \boldsymbol{y}$. Obviously from properties given in §**8.1** the "cross" is a linear operator which, operating on a pair of vectors, yields a vector. We have

$$\boldsymbol{e}_\alpha \times \boldsymbol{e}_\alpha = \boldsymbol{0},\ \boldsymbol{e}_\alpha \times \boldsymbol{e}_\beta = -\boldsymbol{e}_\beta \times \boldsymbol{e}_\alpha.$$

Hence

$$
\begin{aligned}
\boldsymbol{x} \times \boldsymbol{y} = (x_1y_2 - x_2y_1)(\boldsymbol{e}_1 \times \boldsymbol{e}_2) &+ (x_2y_3 - x_3y_2)(\boldsymbol{e}_2 \times \boldsymbol{e}_3)\\
&+ (x_3y_1 - x_1y_3)(\boldsymbol{e}_3 \times \boldsymbol{e}_1),
\end{aligned}
$$

which may be written in the form of a determinant,

$$\boldsymbol{x} \times \boldsymbol{y} = \begin{vmatrix} x_1 & y_1 & (\boldsymbol{e}_2 \times \boldsymbol{e}_3)\\ x_2 & y_2 & (\boldsymbol{e}_3 \times \boldsymbol{e}_1)\\ x_3 & y_3 & (\boldsymbol{e}_1 \times \boldsymbol{e}_2) \end{vmatrix}.$$

Exercises

8.1. If $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$ constitute a basis for a linear vector space, then the vectors $\boldsymbol{e}_2 \times \boldsymbol{e}_3$, $\boldsymbol{e}_3 \times \boldsymbol{e}_1$, $\boldsymbol{e}_1 \times \boldsymbol{e}_2$ form a basis for the same space.

8.2. Show that $(\boldsymbol{x} + \boldsymbol{y}) \times (\boldsymbol{z} + \boldsymbol{w}) = (\boldsymbol{x} \times \boldsymbol{z}) + (\boldsymbol{x} \times \boldsymbol{w}) + (\boldsymbol{y} \times \boldsymbol{z}) + (\boldsymbol{y} \times \boldsymbol{w})$.

8.3. A necessary and sufficient condition that two vectors be linearly dependent is that their vector cross product be the zero-vector.

8.4. Obtain an equation of the line which passes through a point A whose position vector is $\boldsymbol{a}$, and which is perpendicular to two linearly independent vectors $\boldsymbol{b}$ and $\boldsymbol{c}$.

8.5. Verify the following properties of the base vectors $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ constituting a right-handed unitary orthogonal system:

$$\begin{aligned} \boldsymbol{i} \cdot \boldsymbol{i} &= \boldsymbol{j} \cdot \boldsymbol{j} = \boldsymbol{k} \cdot \boldsymbol{k} = 1 \\ \boldsymbol{i} \cdot \boldsymbol{j} &= \boldsymbol{j} \cdot \boldsymbol{k} = \boldsymbol{k} \cdot \boldsymbol{i} = 0 \\ \boldsymbol{i} \times \boldsymbol{i} &= \boldsymbol{j} \times \boldsymbol{j} = \boldsymbol{k} \times \boldsymbol{k} = \boldsymbol{0} \\ \boldsymbol{i} \times \boldsymbol{j} &= \boldsymbol{k}, \boldsymbol{j} \times \boldsymbol{k} = \boldsymbol{i}, \boldsymbol{k} \times \boldsymbol{i} = \boldsymbol{j}. \end{aligned}$$

8.6. If $\boldsymbol{x}$ and $\boldsymbol{y}$ are expressed in terms of an $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ system,

$$\boldsymbol{x} \times \boldsymbol{y} = \begin{vmatrix} x_1 & y_1 & \boldsymbol{i} \\ x_2 & y_2 & \boldsymbol{j} \\ x_3 & y_3 & \boldsymbol{k} \end{vmatrix}.$$

8.7. Forming the scalar product of two vectors is a linear operation, which, performed on two vectors, yields a scalar.

8.8. Differentiation and integration of scalar functions are linear operations.

§9. Products Involving More Than Two Vectors

In this section we develop some of the more important formulas involving scalar and vector products arising from combinations of three or more vectors.

9.1 The scalar triple product $\boldsymbol{z} \cdot (\boldsymbol{x} \times \boldsymbol{y})$.

Consider $\boldsymbol{z} \cdot (\boldsymbol{x} \times \boldsymbol{y})$, where $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}$ are any three vectors. Since $\boldsymbol{x} \times \boldsymbol{y}$ is orthogonal to $\boldsymbol{x}$ and to $\boldsymbol{y}$, and since the scalar product is symmetric,

$$\begin{aligned} \boldsymbol{z} \cdot (\boldsymbol{x} \times \boldsymbol{y}) &= (\boldsymbol{x} \times \boldsymbol{y}) \cdot \boldsymbol{z} \\ \boldsymbol{x} \cdot (\boldsymbol{x} \times \boldsymbol{y}) &= \boldsymbol{y} \cdot (\boldsymbol{x} \times \boldsymbol{y}) = 0 \\ (\lambda \boldsymbol{z}) \cdot (\boldsymbol{x} \times \boldsymbol{y}) &= \lambda\{\boldsymbol{z} \cdot (\boldsymbol{x} \times \boldsymbol{y})\}. \end{aligned}$$

Supposing that the three vectors are linearly independent,

let us consider the parallelepiped whose edges are x, y, and z (Fig. 16).

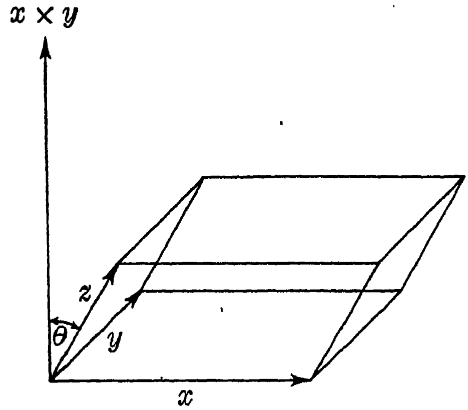

Fig. 16

Then $x \times y$ is a vector which is perpendicular to a face of the parallelepiped whose sides are x and y, and whose length is equal to the area of the face. The scalar product $z \cdot (x \times y)$ is the product of the length of $x \times y$ and the orthogonal projection of z upon it. But this orthogonal projection is simply the altitude of the parallelepiped with a face whose sides are x and y taken as the base of the parallelepiped. That is, $z \cdot (x \times y)$ is a scalar whose numerical value is the *volume* of the parallelepiped whose edges are x, y, and z. By taking different faces in turn as the base of the parallelepiped, we compute its volume in three ways, the results being

$$\text{Volume} = z \cdot (x \times y) = x \cdot (y \times z) = y \cdot (z \times x).$$

Since $y \times x = -(x \times y)$, we also have

$$\text{Volume} = -z \cdot (y \times x) = -x \cdot (z \times y) = -y \cdot (x \times z).$$

Because of this important interpretation we shall call the scalar $z \cdot (x \times y)$ the "box product," and we denote it by the symbol $[xyz]$. In this notation we have

$$[xyz] = [yzx] = [zxy] = -[zyx] = -[xzy] = -[yxz].$$

That is, a cyclic permutation on the letters x, y, and z leaves the value of the box product unchanged; any other permutation merely changes the algebraic sign of the box product.

Let x, y, and z be expressed in terms of a basis

$$\begin{aligned} x &= x_1e_1 + x_2e_2 + x_3e_3 \\ y &= y_1e_1 + y_2e_2 + y_3e_3 \\ z &= z_1e_1 + z_2e_2 + z_3e_3. \end{aligned}$$

Then

$$[zxy] = (z_1e_1 + z_2e_2 + z_3e_3) \cdot \begin{vmatrix} x_1 & y_1 & (e_2 \times e_3) \\ x_2 & y_2 & (e_3 \times e_1) \\ x_3 & y_3 & (e_1 \times e_2) \end{vmatrix}$$

$$= \begin{vmatrix} x_1 & y_1 & z_1[e_1e_2e_3] \\ x_2 & y_2 & 0 \\ x_3 & y_3 & 0 \end{vmatrix} + \begin{vmatrix} x_1 & y_1 & 0 \\ x_2 & y_2 & z_2[e_1e_2e_3] \\ x_3 & y_3 & 0 \end{vmatrix} + \begin{vmatrix} x_1 & y_1 & 0 \\ x_2 & y_2 & 0 \\ x_3 & y_3 & z_3[e_1e_2e_3] \end{vmatrix}.$$

Since

$$[xyz] = [zxy],$$

we may write

$$[xyz] = \begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix} [e_1e_2e_3].$$

9.2 The vector triple product $z \times (x \times y)$.

The vector which arises from the vector triple product $z \times (x \times y)$, where x, y, and z are linearly independent, is a vector which is parallel to the plane determined by x and y. Hence we may write

$$z \times (x \times y) = \lambda x + \mu y,$$

where λ and μ are scalars to be determined. Operating on each side of this equation by $z\cdot$, we have

$$0 = \lambda z \cdot x + \mu z \cdot y.$$

Hence, λ and μ may be written in the forms

$$\lambda = \nu z \cdot y, \mu = -\nu z \cdot x,$$

where ν is a scalar to be determined. We now have

$$z \times (x \times y) = \nu\{(z \cdot y)x - (z \cdot x)y\}.$$

If we operate on each side of this equation by $x \cdot$, it results that

$$\{z \times (x \times y)\} \cdot x = \nu\{(z \cdot y)(x \cdot x) - (z \cdot x)(y \cdot x)\}.$$

This equation will give the value of ν if we can evaluate the left member. We may write

$$\{z \times (x \times y)\} \cdot x = \{(x \times y) \times x\} \cdot z = -\{x \times (x \times y)\} \cdot z.$$

Consider then $x \times (x \times y)$, which is a vector in the plane of x and y, and which is perpendicular to x. The length of

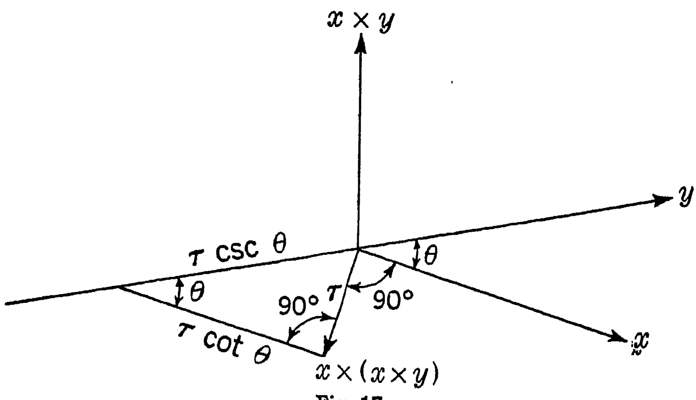

Fig. 17

this vector, which we denote by τ, is seen to be the numerical value of $|x|^2|y| \sin\theta$, where θ is the angle between x and y. The components of $x \times (x \times y)$ on x and $-y$ are $\tau \cot\theta$ and $\tau \csc\theta$, respectively, or $|x|^2|y| \cos\theta$ and $|x|^2|y|$. Hence

$$\begin{aligned} x \times (x \times y) &= (|x|^2|y| \cos\theta)\frac{x}{|x|} - (|x|^2|y|)\frac{y}{|y|} \\ &= (x \cdot y)x - (x \cdot x)y. \end{aligned}$$

When we make use of this result, it follows that the scalar ν

in the above equation must have the value $+1$. We then have

$$z \times (x \times y) = (z \cdot y)x - (z \cdot x)y = \begin{vmatrix} x & y \\ z \cdot x & z \cdot y \end{vmatrix},$$

a fundamental relation.

9.3 The Lagrange identity.

We now establish the important formula

$$(x \times y) \cdot (\bar{x} \times \bar{y}) = \begin{vmatrix} x \cdot \bar{x} & x \cdot \bar{y} \\ y \cdot \bar{x} & y \cdot \bar{y} \end{vmatrix},$$

frequently called the *Lagrange identity*.

Set $w = \bar{x} \times \bar{y}$; then

$$\begin{aligned} (x \times y) \cdot (\bar{x} \times \bar{y}) &= (x \times y) \cdot w = (y \times w) \cdot x \\ &= \{y \times (\bar{x} \times \bar{y})\} \cdot x \\ &= \begin{vmatrix} \bar{x} & \bar{y} \\ y \cdot \bar{x} & y \cdot \bar{y} \end{vmatrix} \cdot x \\ &= \begin{vmatrix} x \cdot \bar{x} & x \cdot \bar{y} \\ y \cdot \bar{x} & y \cdot \bar{y} \end{vmatrix}. \end{aligned}$$

9.4 Reciprocal system of vectors.

References: *Wills* (23), p. 39; *Gibbs-Wilson* (7), p. 81; *Runge* (17), p. 46; *Lagally* (13), p. 37.

Let e_1, e_2, e_3 form a basis. We now introduce three vectors e^1, e^2, e^3 defined by

$$(A) \qquad e_\alpha \cdot e^\beta = \begin{cases} 0 \text{ for } \alpha \neq \beta \\ 1 \text{ for } \alpha = \beta \end{cases}.$$

The vectors e^1, e^2, e^3 are thus perpendicular to the planes determined by e_2, e_3; e_3, e_1; e_1, e_2, respectively. Hence

$$e^1 = \alpha(e_2 \times e_3),$$

where α is as yet an undetermined scalar. But

$$e_1 \cdot e^1 = 1 = \alpha e_1 \cdot (e_2 \times e_3) = \alpha[e_1 e_2 e_3].$$

Since e_1, e_2, e_3 are linearly independent, $[e_1 e_2 e_3] \neq 0$ (see

Exercise 9.2), and therefore

$$\alpha = \frac{1}{[e_1e_2e_3]}.$$

We thus have

$$e^1 = \frac{e_2 \times e_3}{[e_1e_2e_3]}, \quad e^2 = \frac{e_3 \times e_1}{[e_1e_2e_3]}, \quad e^3 = \frac{e_1 \times e_2}{[e_1e_2e_3]}.$$

The vectors e^1, e^2, e^3 constitute a basis for the vector space under consideration. For, let

$$\lambda_1 e^1 + \lambda_2 e^2 + \lambda_3 e^3 = 0.$$

Operating on each side of this equation with $e_1 \cdot$, we obtain $\lambda_1 = 0$. Similarly $\lambda_2 = 0$ and $\lambda_3 = 0$. Thus the vectors are linearly independent.

Two sets of base vectors satisfying the conditions (A) are said to be *reciprocal systems* with respect to one another.

Exercises

9.1. From a consideration of $(x \times y) \cdot (x \times y)$, establish the Cauchy-Schwarz inequality.

9.2. A necessary and sufficient condition that three vectors be linearly dependent is that their box product shall vanish.

9.3. Given three linearly independent vectors a, b, c, show that the unique solution of the system

$$a \cdot x = \alpha, \; b \cdot x = \beta, \; c \cdot x = \gamma,$$

where α, β, γ are any three real numbers, is given by

$$x = \frac{\alpha(b \times c) + \beta(c \times a) + \gamma(a \times b)}{[abc]}.$$

9.4. Establish the formulas

$$\begin{aligned}(a \times b) \times (c \times d) &= [acd]b - [bcd]a \\ &= [abd]c - [abc]d.\end{aligned}$$

For an interpretation of these formulas in connection with spherical trigonometry, see *Gibbs-Wilson* (7), p. 77.

9.5. If e_1, e_2, e_3 constitute a basis, and a vector x is expressed in the form

$$x = \alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3,$$

the coefficients have the values

$$\alpha_1 = \frac{[\boldsymbol{x}\boldsymbol{e}_2\boldsymbol{e}_3]}{[\boldsymbol{e}_1\boldsymbol{e}_2\boldsymbol{e}_3]}, \ \alpha_2 = \frac{[\boldsymbol{e}_1\boldsymbol{x}\boldsymbol{e}_3]}{[\boldsymbol{e}_1\boldsymbol{e}_2\boldsymbol{e}_3]}, \ \alpha_3 = \frac{[\boldsymbol{e}_1\boldsymbol{e}_2\boldsymbol{x}]}{[\boldsymbol{e}_1\boldsymbol{e}_2\boldsymbol{e}_3]},$$

which may also be written

$$\alpha_1 = \boldsymbol{x} \cdot \boldsymbol{e}^1, \ \alpha_2 = \boldsymbol{x} \cdot \boldsymbol{e}^2, \ \alpha_3 = \boldsymbol{x} \cdot \boldsymbol{e}^3,$$

where $\boldsymbol{e}^1$, $\boldsymbol{e}^2$, $\boldsymbol{e}^3$ form the reciprocal basis.

9.6. Let $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$ and $\boldsymbol{e}^1$, $\boldsymbol{e}^2$, $\boldsymbol{e}^3$ be reciprocal systems. Establish the following:

(1) The reciprocal system of the reciprocal system is the original system;

(2) $[\boldsymbol{e}_1\boldsymbol{e}_2\boldsymbol{e}_3]\,[\boldsymbol{e}^1\boldsymbol{e}^2\boldsymbol{e}^3] = 1$;

(3) The only self-reciprocal system of base vectors is an $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ system.

9.5 Covariant and contravariant vectors.

Given a basis $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$ for a linear vector space, we have seen that any vector $\boldsymbol{x}$ of the space can be expressed uniquely in the form

$$\boldsymbol{x} = x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + x_3\boldsymbol{e}_3.$$

Also, the given base vectors determine uniquely the reciprocal system $\boldsymbol{e}^1$, $\boldsymbol{e}^2$, $\boldsymbol{e}^3$, and these in turn constitute a basis for the same space. Hence the same vector $\boldsymbol{x}$ has a unique representation in the form

$$\boldsymbol{x} = \bar{x}_1\boldsymbol{e}^1 + \bar{x}_2\boldsymbol{e}^2 + \bar{x}_3\boldsymbol{e}^3.$$

Usually these two representations will be different; that is, generally the coefficients x_1, x_2, x_3 will be different from $\bar{x}_1$, $\bar{x}_2$, $\bar{x}_3$. If a vector is expressed in terms of the base vectors $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$, it is called a *contravariant vector;* if a vector is expressed in terms of the reciprocal basis $\boldsymbol{e}^1$, $\boldsymbol{e}^2$, $\boldsymbol{e}^3$, it is called a *covariant vector*. Notice that the terms "covariant" and "contravariant" applied to a vector do not characterize a vector but merely the way in which it is described.

It is evident that, when either description of a vector is given, the other description is completely determined. Actually the passage from one description to the other is not very complicated, as the following exercises will reveal. We see that the fundamental quadratic form or, more specifically, its coefficients, will necessarily enter into the change from one description of a vector to the other. For it is only by means of the quadratic form that we are able to introduce the reciprocal system as this system involves in its definition the notions of angle and length. So long as we had no metric available, we could consider vectors only in their contravariant description.

Even though a vector may be described in two ways in a given frame of reference, its *length*, and the *angle* between two vectors, are *numbers independent of the mode of description*. We continue to use the symbols $\boldsymbol{x} \cdot \boldsymbol{x}$, and $\boldsymbol{x} \cdot \boldsymbol{y}$ for the square of the length of $\boldsymbol{x}$, and the scalar product of $\boldsymbol{x}$ and $\boldsymbol{y}$, respectively. However, the method of computing these numbers does depend upon the contravariant or covariant description of the vectors involved.

Exercises

9.7. Let $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ constitute a basis, and let $\lambda_{\mu\nu}$ be the coefficients of the fundamental quadratic form (§**6.1**).

(1) Let $\boldsymbol{x}$ be a vector having the *contravariant* description $\boldsymbol{x} = x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + x_3\boldsymbol{e}_3$ and the *covariant* description $\boldsymbol{x} = \bar{x}_1\boldsymbol{e}^1 + \bar{x}_2\boldsymbol{e}^2 + \bar{x}_3\boldsymbol{e}^3$. Then

$$\bar{x}_1 = \boldsymbol{x} \cdot \boldsymbol{e}_1,\ \bar{x}_2 = \boldsymbol{x} \cdot \boldsymbol{e}_2,\ \bar{x}_3 = \boldsymbol{x} \cdot \boldsymbol{e}_3,$$

and hence

$$\begin{aligned}
\bar{x}_1 &= \lambda_{11}x_1 + \lambda_{12}x_2 + \lambda_{13}x_3\\
\bar{x}_2 &= \lambda_{21}x_1 + \lambda_{22}x_2 + \lambda_{23}x_3\\
\bar{x}_3 &= \lambda_{31}x_1 + \lambda_{32}x_2 + \lambda_{33}x_3.
\end{aligned}$$

(2) In particular, the covariant descriptions of the base vectors $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ are given by

$$\begin{aligned}
\boldsymbol{e}_1 &= \lambda_{11}\boldsymbol{e}^1 + \lambda_{21}\boldsymbol{e}^2 + \lambda_{31}\boldsymbol{e}^3\\
\boldsymbol{e}_2 &= \lambda_{12}\boldsymbol{e}^1 + \lambda_{22}\boldsymbol{e}^2 + \lambda_{32}\boldsymbol{e}^3\\
\boldsymbol{e}_3 &= \lambda_{13}\boldsymbol{e}^1 + \lambda_{23}\boldsymbol{e}^2 + \lambda_{33}\boldsymbol{e}^3.
\end{aligned}$$

(3) Solving the last system of equations, obtain the contravariant descriptions of the vectors e^1, e^2, e^3 in the form

$$\begin{aligned} e^1 &= \lambda^{11}e_1 + \lambda^{12}e_2 + \lambda^{13}e_3 \\ e^2 &= \lambda^{21}e_1 + \lambda^{22}e_2 + \lambda^{23}e_3 \\ e^3 &= \lambda^{31}e_1 + \lambda^{32}e_2 + \lambda^{33}e_3, \end{aligned}$$

where $\lambda^{\mu\nu}$ stands for the cofactor of $\lambda_{\mu\nu}$ in the determinant $|\lambda_{\mu\nu}|$ divided by the determinant.

(4) Show that

$$e^\mu \cdot e^\nu = \lambda^{\mu\nu}.$$

(5) If y is the vector

$$y = y_1e_1 + y_2e_2 + y_3e_3 = \bar{y}_1e^1 + \bar{y}_2e^2 + \bar{y}_3e^3,$$

the scalar product $x \cdot y$ is given by

$\Sigma\lambda_{\alpha\beta}x_\alpha y_\beta$ where x and y are each contravariant;
$\Sigma x_\alpha\bar{y}_\alpha$ or $\Sigma\bar{x}_\alpha y_\alpha$ where one is covariant and the other is contravariant;
$\Sigma\lambda^{\alpha\beta}\bar{x}_\alpha\bar{y}_\beta$ where each is covariant.

(6) The box product $[e_1e_2e_3] = \sqrt{g}$, where g is the determinant $|\lambda_{\alpha\beta}|$.

(7) Obtain the covariant descriptions of the vectors x and y of Exercise **6.12** and verify (5) of this exercise.

9.8. The covariant and contravariant descriptions of an arbitrary vector are the same, if and only if the basis is an i, j, k system.

§10. Applications of the Algebra of Vectors

10.1 Algebra of vectors.

References: *Dickson* (30); *Hardy* (36); *Kelland-Tait* (39).

The material of the preceding sections is spoken of as the "algebra of vectors." However, the vectors do not constitute a division algebra (cf. *Dickson* (35), pp. 59–64), but the operations thus far considered are sufficiently algebraic in character to make the phrase "algebra of vectors" seem appropriate. Actually, the efforts of Hamilton and his co-workers to construct a genuine algebra of hypercomplex numbers resulted in the development of

quaternions. The system of real quaternions comprises an algebra based on four units.

10.2 Moment of a vector.

References: *Gibbs-Wilson* (7), p. 92; *Wills* (23), p. 28; *Jeans* (37), p. 60.

Let x be the position vector of a point P with respect to O, and let y represent a force *acting at* P. If O and P are points of a rigid body with O fixed, the force y tends to produce a rotation of the body about an axis through O which is perpendicular to the plane of the vectors x and y. In mechanics the *moment* of the force y about O is defined as the product of the length of x and the component of y which is perpendicular to x. Thus the moment is simply the area of the parallelogram whose sides are x and y. The vector

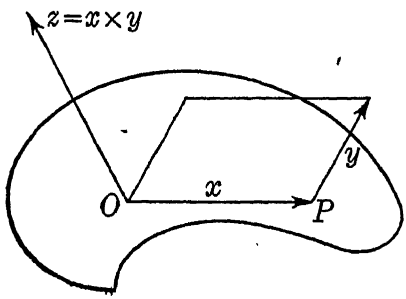

Fig. 18

$$z = x \times y$$

is called the *moment vector* of y with respect to the point O. Clearly, so far as moment is concerned the result is the same if the force y acts not at P but at any point P' along its line of action.

10.3 Couple.

References: *Jeans* (37), p. 99; *Love* (42), p. 200.

Let P, Q be points of a rigid body at which forces $-y$ and y, respectively, are acting. This system of forces is called a *couple.* The moment vector of the couple about P is given by $\overrightarrow{PQ} \times y$. Let O be any point and let a, b be the position vectors of P, Q, respectively, with reference to O. Then the moment vector z of the couple about O is given by

$$\begin{aligned} z &= -(a \times y) + (b \times y) \\ &= (b - a) \times y \\ &= \overrightarrow{PQ} \times y. \end{aligned}$$

Thus the moment vector of a couple is the same with respect to every point.

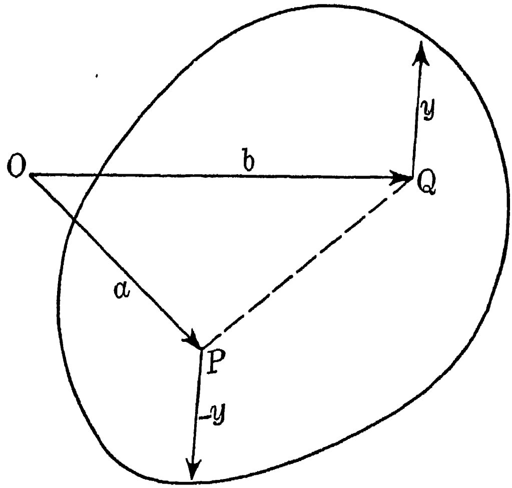

Fig. 19

10.4 Motion of a rigid body.

References: *Jeans* (37), pp. 106–107; *Wills* (23), p. 34.

Suppose forces denoted by the vectors $f_1, f_2, \ldots, f_n$ act at points $P_1, P_2, \ldots, P_n$, respectively, of a rigid body. Let O be any given point of the rigid body, and denote the position vectors of $P_1, P_2, \ldots, P_n$ from O by $a_1, a_2, \ldots, a_n$, respectively. We now make the following assumptions: (1) the effect of any number of forces acting at a point of a body is the same as that of their resultant acting at that point; (2) the effect of any number of couples acting on a rigid body is the same as that produced by a resultant couple, the moment vector of which is obtained by adding the individual moment vectors of the various couples (cf. *Love* (42), p. 21 and p. 203).

With these hypotheses we now prove that the system of forces acting on a rigid body as above described is equiva-

lent to a single force acting at an arbitrarily assigned point in the body, combined with a couple.

At a point O in the body let forces f_1 and $-f_1$ be introduced. There is then a force f_1 at O and the couple determined by f_1 at P_1 and $-f_1$ at O. Let this procedure be

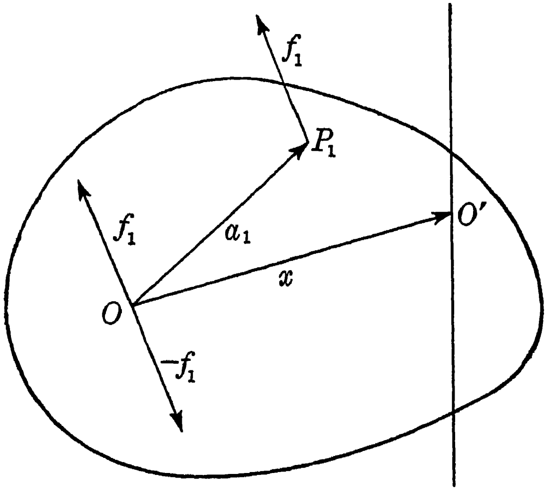

Fig. 20

followed for each of the forces $f_2, f_3, \ldots, f_n$. Thus we have a resultant force f at O given by

$$f = f_1 + f_2 + \cdots + f_n$$

and a couple whose moment vector g is given by

$$g = (a_1 \times f_1) + (a_2 \times f_2) + \cdots + (a_n \times f_n).$$

These relations hold for any point O in the rigid body. Let us now select O, if possible, so that the vectors f and g will be parallel. Let O' be such a point, and let its position vector with respect to O be x. Let the position vectors of $P_1, P_2, \ldots, P_n$ with respect to O' be $a_1', a_2', \ldots, a_n'$, respectively. Then

$$a_1' = a_1 - x,\ a_2' = a_2 - x,\ \cdots,\ a_n' = a_n - x.$$

Hence g', the resultant moment vector, is given by

$$g' = (a_1 - x) \times f_1 + (a_2 - x) \times f_2 + \cdots + (a_n - x) \times f_n;$$

that is

$$g' = g - (x \times f).$$

The object is to determine x, if possible, so that g' and f shall be parallel. This condition being imposed, x must satisfy the equation

$$x \times f = g - \mu f,$$

where μ is a scalar. We note that if $\bar{x}$ is a solution, then

$$x = \bar{x} + \lambda f,$$

where λ is an arbitrary scalar, is also a solution. Hence, if there is a solution, the point O' can be any point on a certain line which is parallel to f. This line is called the *central axis* of the system of forces.

Since

$$f \cdot (x \times f) = 0$$

we have

$$f \cdot g - \mu f \cdot f = 0;$$

that is,

$$\mu = \frac{f \cdot g}{f \cdot f} \text{ (we suppose } f \neq 0\text{)}.$$

We now have

$$x \times f = g - \frac{f \cdot g}{f \cdot f} f,$$

and the problem has been reduced to that of finding a particular solution x of this equation. Writing the right member in the form

$$\frac{1}{f \cdot f}\begin{vmatrix} g & f \\ g \cdot f & f \cdot f \end{vmatrix},$$

we recognize it as the expansion of

$$\frac{f \times (g \times f)}{f \cdot f}.$$

Hence a particular solution of the above equation is

$$\bar{x} = -\frac{g \times f}{f \cdot f}.$$

We have thus determined the central axis of the system of forces; the resultant force f acts along this line, and the moment vector g' of the resultant couple is parallel to this line. Since any couple producing the moment vector g' consists of a pair of vectors in a plane perpendicular to g', such a couple tends to produce a rotation about the central axis. Thus, any system of forces acting upon a rigid body is equivalent to a single force which tends to produce a translation along the central axis, and a couple which tends to produce a rotation about this axis.

10.5 Angular velocity vector.

Suppose a rigid body is rotating about a fixed axis l with a constant velocity of τ radians per second. The rotation is completely described by a vector z, known as the *angular velocity vector*,[9] which has the properties

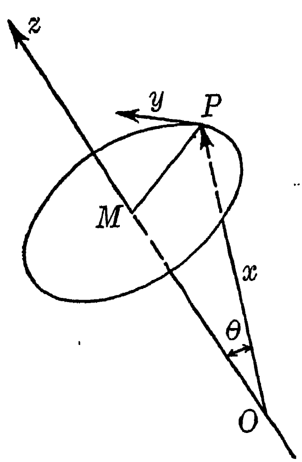

Fig. 21

(1) z is parallel to the axis of rotation l;

(2) z has for its length the number τ;

(3) z has as its positive direction that in which a right-handed screw would advance under the given rotation.

The rotation about l imparts to each point P of the body a *linear velocity* which we denote by the vector y. Let O be any point on the axis l, and let $\overrightarrow{OP} = x$. Since P describes a circle with center M, a point on the axis of rotation, the linear velocity vector y of P is perpendicular to MP and has as its length the value $\tau\overline{MP}$. If θ is the angle between x and z, then

[9] Cf. p. 84.

$$\overline{MP} = |x| \sin \theta.$$

But r is the length of z. Therefore, $y = z \times x$; that is,

(linear velocity vector) = (angular velocity vector) $\times$ (position vector).

10.6 Finite rotation about a line not a vector.

This is perhaps a suitable time to point out that something more is required of a vector quantity than merely that it is a quantity which can be described by a directed line segment. Consider a finite rotation of a rigid body about a fixed axis. This can obviously be unambiguously described by a directed line segment taken along the axis of rotation with proper sense of direction and a length which measures the angle of rotation. These directed line segments, however, do not satisfy the commutative law of addition, and hence are not vectors. A particular case will suffice to show this.

Let e_1 and e_2 be two non-orthogonal unit vectors which are linearly independent and have an initial point O in

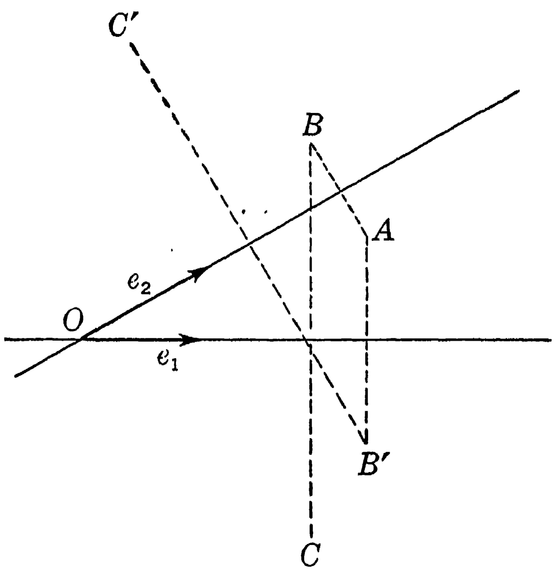

Fig. 22

common. Let A be a point in the plane determined by e_1, e_2. Consider a rotation of A through an angle π about e_2 followed by an equal rotation about e_1. These rotations in this order displace A into a point C; if the rotations are carried out in the other order, A is displaced into a point C' different from C. Thus a finite rotation about a line is not a vector quantity. For a discussion of orthogonal line reflections, reference may be made to *Veblen-Young* (51), II.

Care is needed in the application of any mathematical subject to particular instances. An application usually involves assumptions additional to those made in the mathematics. For instance, in the abstract development of the algebra of vectors a parallel displacement was without significance, but this is clearly not the case when we consider forces acting upon a rigid body where the points at which the forces act play an essential part.

The subject of Graphical Statics may be regarded largely as an instance of the algebra of vectors. *Carathéodory* (29), Chapter VI, makes use of vectors in the development of determinants, and in problems of intersection of linear spaces. *Schreier-Sperner* (49), II, make extensive use of the vector analysis of an $(n + 1)$-affine space in their treatment of the projective geometry of an n-space.

Exercises

10.1. Let 1, i, j, k be symbols having the following *multiplication* table:

$\nearrow$	1	i	j	k
1	1	i	j	k
i	i	-1	k	$-j$
j	j	$-k$	-1	i
k	k	j	$-i$	-1

Then an expression of the form

$$q = \delta + \alpha i + \beta j + \gamma k,$$

where the coefficients δ, α, β, γ are real numbers is called a real

quaternion. We write

$$q = Sq + Vq,$$

where

$Sq = \delta$ is called the "scalar part" of the quaternion;
$Vq = \alpha i + \beta j + \gamma k$ is called the "vector part" of the quaternion.

The sum and product of two quaternions

$$\begin{aligned} q_1 &= \delta_1 + \alpha_1 i + \beta_1 j + \gamma_1 k \\ q_2 &= \delta_2 + \alpha_2 i + \beta_2 j + \gamma_2 k \end{aligned}$$

are formed as the sum and product of two linear polynomials in algebra, and in the case of a product the result is reduced by means of the above multiplication table.

(1) Show that the sum and product of any two quaternions is a quaternion.

(2) Obtain the product q_1q_2 and show that it is not in general equal to q_2q_1.

(3) A vector may be regarded as a quaternion whose scalar part is zero. Let

$$\begin{aligned} a_1 &= \alpha_1 i + \beta_1 j + \gamma_1 k \\ a_2 &= \alpha_2 i + \beta_2 j + \gamma_2 k \end{aligned}$$

be two given vectors. Regarding these vectors as quaternions, obtain their product and show that

$$\begin{aligned} S(a_1a_2) &= S(a_2a_1) = -a_1 \cdot a_2 \\ V(a_1a_2) &= -V(a_2a_1) = a_1 \times a_2. \end{aligned}$$

(4) If q is the quaternion

$$q = Sq + Vq,$$

the *conjugate* of q, which we denote by $\bar{q}$, is defined by

$$\bar{q} = Sq - Vq.$$

Prove the following:

(*a*) Forming the conjugate of a quaternion is a linear operation of period 2.

(*b*) Negation, reciprocation, and conjugation applied to quaternions are commutative processes.

(*c*) $Sq = \frac{1}{2}(q + \bar{q})$, $Vq = \frac{1}{2}(q - \bar{q})$.

(*d*) $\overline{q_1q_2} = \bar{q}_2\bar{q}_1$.

(*e*) $q\bar{q} = \bar{q}q$.

(5) The product $q\bar{q}$ is called the *norm* of q, which we denote by Nq. Prove the following:

(*a*) $Nq = N\bar{q}$.

(*b*) Nq is a non-negative scalar.

(*c*) $Nq = 0$ is a necessary and sufficient condition that the real quaternion q be the *zero* quaternion, that is, the quaternion all of whose coefficients are zero.

(6) Prove

$$S(q_1q_2) = \tfrac{1}{2}(q_1q_2 + \bar{q}_2\bar{q}_1)$$
$$V(q_1q_2) = \tfrac{1}{2}(q_1q_2 - \bar{q}_2\bar{q}_1).$$

(7) Solve the equation

$$qq_1 = q_2,$$

where

$$q_1 = 3 + 2i - j + k.$$
$$q_2 = 2 - i + j + k.$$

10.2. Show that a system of forces represented in magnitude, direction, and position by the sides of a plane polygon all directed clockwise (or anticlockwise) is equivalent to a couple whose moment is equal to twice the area of the polygon.

10.3. Show that any system of forces acting on a rigid body can be reduced to two equal forces equally inclined to the central axis.

10.4. Obtain the equation of the line which passes through a given point and which intersects two given lines.

10.5. Obtain the shortest distance between two given lines.

10.6. Let $\boldsymbol{a}$ and $\boldsymbol{b}$ be unit vectors in the $\boldsymbol{i}$, $\boldsymbol{j}$ plane making angles α, β, respectively, with the vector $\boldsymbol{i}$. By considering $\boldsymbol{a} \cdot \boldsymbol{b}$ and $\boldsymbol{a} \times \boldsymbol{b}$, deduce the trigonometric formulas for cos $(\alpha - \beta)$ and sin $(\alpha - \beta)$.

10.7. Show that

$$[\boldsymbol{xyz}][\boldsymbol{abc}] = \begin{vmatrix} \boldsymbol{x} \cdot \boldsymbol{a} & \boldsymbol{x} \cdot \boldsymbol{b} & \boldsymbol{x} \cdot \boldsymbol{c} \\ \boldsymbol{y} \cdot \boldsymbol{a} & \boldsymbol{y} \cdot \boldsymbol{b} & \boldsymbol{y} \cdot \boldsymbol{c} \\ \boldsymbol{z} \cdot \boldsymbol{a} & \boldsymbol{z} \cdot \boldsymbol{b} & \boldsymbol{z} \cdot \boldsymbol{c} \end{vmatrix}.$$

Note that this may be considered as a theorem in the multiplication of determinants.

10.8. Show that

$$\boldsymbol{x} \times (\boldsymbol{y} \times \boldsymbol{z}) + \boldsymbol{y} \times (\boldsymbol{z} \times \boldsymbol{x}) + \boldsymbol{z} \times (\boldsymbol{x} \times \boldsymbol{y}) = \boldsymbol{0}.$$

10.9. The bisectors of the interior angles of a triangle meet in a point.

10.10. Explain why a sphere rolling down an inclined plane does not contradict the theorem deduced in §**10.4.**

10.11. Let points $P_1, P_2, \ldots, P_n$ be endowed with numbers (or masses) $\lambda_1, \lambda_2, \ldots, \lambda_n$, respectively, where λ_i are real numbers whose sum is not zero. Let Q be any point in space:

(1) The resultant of the forces $\lambda_1\overrightarrow{QP_1}, \lambda_2\overrightarrow{QP_2}, \ldots, \lambda_n\overrightarrow{QP_n}$, applied at Q, passes through a fixed point C and is equal to $\lambda\overrightarrow{QC}$, where λ denotes the sum $\lambda_1 + \lambda_2 + \cdots + \lambda_n$.

(2) The position vector of the point C is

$$\overrightarrow{OC} = \frac{\sum_i \lambda_i\overrightarrow{OP_i}}{\lambda}.$$

(3) In the case of three points, the line P_1C divides the line segment $\overrightarrow{P_2P_3}$ in the ratio $\lambda_3:\lambda_2$.

(4) Consider (1) and (2) where λ_i is replaced by λ_i/k, k being the length of the vector $\overrightarrow{QP_1}$, and let Q recede indefinitely along any straight line. Thus obtain by a limiting process the usual theorem concerning the center of gravity of n point masses subject to parallel forces which are proportional to their masses.

CHAPTER II

Differential Calculus of Vectors

§11. Vector Function of a Scalar

References: *Gibbs-Wilson* (7), Chapter III; *Juvet* (11), pp. 33–35; *Phillips* (15), pp. 30–32.

11.1 Variable vector as function of a scalar.

We now consider a variable vector $\boldsymbol{x}$ which depends upon a scalar variable t. The vector $\boldsymbol{x}(t)$ is said to be a *function* of the real scalar variable t on the interval $t_1 \leqq t \leqq t_2$, if to each value of t on the interval there corresponds the unique vector $\boldsymbol{x}(t)$. An instance would be the position vector of a moving particle regarded as a function of the time, or the velocity vector of such a particle.

11.2 Limit of a vector.

Given $\boldsymbol{x}(t)$ defined on $[t_1, t_2]$, where we write $[t_1, t_2]$ for the interval $t_1 \leqq t \leqq t_2$, let us consider what we shall mean by the "limit of $\boldsymbol{x}(t)$ as t approaches t_0" where t_0 is a value on the interval. We write

$$\boldsymbol{a} = \lim_{t \to t_0} \boldsymbol{x}(t),$$

and by this we mean: that, given any positive number ϵ, there exists a positive number η (which may depend upon ϵ and t_0), such that for all t on $[t_1, t_2]$ satisfying the inequality $0 < |t - t_0| < \eta$, it is true that $|\boldsymbol{a} - \boldsymbol{x}(t)| < \epsilon$. Geometrically this means that, if the vectors are laid off from a point O as initial point, the end points P of the vectors $\overrightarrow{OP} = \boldsymbol{x}(t)$ will be within a sphere of radius ϵ and center at the end point of $\boldsymbol{a}$ for all t satisfying the inequality $0 < |t - t_0| < \eta$. If $\boldsymbol{x}(t)$ and t_0 are such that there is no constant vector $\boldsymbol{a}$ satisfying these conditions, then the limit of $\boldsymbol{x}(t)$ as t approaches t_0 fails to exist.

The vector function $\boldsymbol{x}(t)$ is said to be a *continuous function of t, for $t = t_0$* if the following three conditions hold:

(1) $\boldsymbol{x}(t)$ is defined for $t = t_0$;
(2) $\lim \boldsymbol{x}(t)$ exists as t approaches t_0;
(3) $\lim_{t \to t_0} \boldsymbol{x}(t) = \boldsymbol{x}(t_0)$.

If one or more of the second and third conditions fail to hold at a point, that is, for a value of t, the function $\boldsymbol{x}(t)$ is said to be *discontinuous* at that point. The function $\boldsymbol{x}(t)$ is said to be a continuous function of t on the interval $[t_1, t_2]$ if it is continuous at each point of the interval.

11.3 Differentiation of a vector.

Given $\boldsymbol{x}(t)$, then $\boldsymbol{x}(t + \Delta t) - \boldsymbol{x}(t)$, where Δt denotes an increment in t, is of course a vector. If

$$\lim_{\Delta t \to 0} \frac{\boldsymbol{x}(t + \Delta t) - \boldsymbol{x}(t)}{\Delta t}$$

exists, it is called the *derivative of $\boldsymbol{x}(t)$ with respect to t.* We denote this derived vector by the notation $d\boldsymbol{x}/dt$ or $\dot{\boldsymbol{x}}(t)$. Similarly we designate $d\dot{\boldsymbol{x}}(t)/dt$ by $\ddot{\boldsymbol{x}}(t)$.

Let O; $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$ constitute an affine coördinate system and let the vectors $\boldsymbol{x}(t)$ be laid off from the origin as the initial point. Let $\boldsymbol{x}(t)$ be expressed in terms of the base vectors

$$\boldsymbol{x}(t) = x_1(t)\boldsymbol{e}_1 + x_2(t)\boldsymbol{e}_2 + x_3(t)\boldsymbol{e}_3.$$

Then

$$\boldsymbol{x}(t + \Delta t) = x_1(t + \Delta t)\boldsymbol{e}_1 + x_2(t + \Delta t)\boldsymbol{e}_2 + x_3(t + \Delta t)\boldsymbol{e}_3.$$

Hence

$$\begin{aligned}\frac{d\boldsymbol{x}}{dt} &= \lim_{\Delta t \to 0} \frac{\boldsymbol{x}(t + \Delta t) - \boldsymbol{x}(t)}{\Delta t} \\ &= \lim_{\Delta t \to 0} \left\{ \frac{x_1(t + \Delta t) - x_1(t)}{\Delta t}\boldsymbol{e}_1 + \frac{x_2(t + \Delta t) - x_2(t)}{\Delta t}\boldsymbol{e}_2 \right. \\ &\qquad \left. + \frac{x_3(t + \Delta t) - x_3(t)}{\Delta t}\boldsymbol{e}_3 \right\};\end{aligned}$$

that is,

$$\frac{d\boldsymbol{x}}{dt} = \frac{dx_1}{dt}\boldsymbol{e}_1 + \frac{dx_2}{dt}\boldsymbol{e}_2 + \frac{dx_3}{dt}\boldsymbol{e}_3.$$

It is important to recognize that this result is obtained on the assumption that the base vectors e_1, e_2, e_3 are constant. This would not be the case were a curvilinear coördinate system employed, such as the polar coördinate system in the plane.

Exercises

11.1. The postage on first-class mail is a function of the weight; let it be $p(w)$, where the weight is in ounces. Show that the function $p(w)$ is discontinuous for positive integral values of w.

11.2. If $x(t) = x_1(t)e_1 + x_2(t)e_2 + x_3(t)e_3$ for t on $[t_1, t_2]$, a necessary and sufficient condition that $x(t)$ be continuous on the interval is that the scalar functions $x_1(t)$, $x_2(t)$, and $x_3(t)$ be continuous on the interval.

11.3. If x, y, and z are vector functions of t, establish the following rules of differentiation:

(1) $\frac{d}{dt}(x \pm y) = \frac{dx}{dt} \pm \frac{dy}{dt}$

(2) $\frac{d}{dt}\{\lambda(t)x\} = \frac{d\lambda}{dt}x + \lambda\frac{dx}{dt}$

(3) $\frac{d}{dt}(x \cdot y) = \frac{dx}{dt} \cdot y + x \cdot \frac{dy}{dt}$

(4) $\frac{d}{dt}(x \times y) = \frac{dx}{dt} \times y + x \times \frac{dy}{dt}$

(5) $\frac{d}{dt}[xyz] = \left[\frac{dx}{dt}yz\right] + \left[x\frac{dy}{dt}z\right] + \left[xy\frac{dz}{dt}\right].$

Note that (5) gives a rule for differentiating a determinant.

11.4. If the derivative $\dot{x}(t)$ exists, $x(t)$ is necessarily continuous.

§12. Geometry of Space Curves

References: *Blaschke* (26), Chapter I; *Juvet* (11), Chapter II; *Wills* (23), pp. 56–62.

12.1 Vector equation of a curve.

Let $x(t)$, $t_1 \leqq t \leqq t_2$ be a continuous vector function of t possessing as many derivatives as may be needed in the

discussion. Let O be a fixed point. Then the end points P of the vectors $\overrightarrow{OP} = \mathbf{x}(t)$ constitute a *one-dimensional set of points* which is called a *curve.*

As an illustration of the differential calculus of vectors, we shall consider an introduction to the differential geometry of space curves. This is not as specialized as it may perhaps seem at first. For, no matter what the interpretation of the vector function $\mathbf{x}(t)$ is in a given case, there is associated a curve C, unique except for its position in space, and any geometric property of the curve has significance with respect to the given problem.

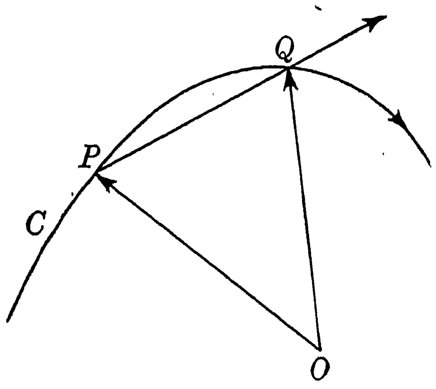

Fig. 23.

Let C (Fig. 23) be the curve defined by

$$C: \quad \mathbf{x} = \mathbf{x}(t), \; t_1 \leqq t \leqq t_2.$$

We suppose that, as t increases continuously on the range $[t_1, t_2]$, the corresponding point P moves continuously along C in a certain direction. We assign this as the positive direction along C.

12.2 Tangent to a curve.

Definition: Let P be a fixed point on C and let Q be a variable point on C (Fig. 23). The line through P and Q is called a *secant line.* If the line PQ approaches a limiting

position as Q approaches P, this limit line is called the *tangent to the curve C at the point P.*

Let P correspond to t and Q to $t + \Delta t$. Then

$$\overrightarrow{OP} = x(t), \ \overrightarrow{OQ} = x(t + \Delta t)$$

and

$$\overrightarrow{PQ} = x(t + \Delta t) - x(t).$$

Also, $\overrightarrow{PQ}/\Delta t$ is a vector parallel to $\overrightarrow{PQ}$. Hence the tangent line at P is parallel to

$$\lim_{\Delta t \to 0} \frac{x(t + \Delta t) - x(t)}{\Delta t} = \frac{dx(t)}{dt} \text{ or } \dot{x}(t).$$

The vector $\dot{x}(t)$ is called the *tangent vector* to the curve C at the point P. We shall assume throughout that the vector $\dot{x}(t) \neq 0$.

The equation of the tangent line is

$$y = x + \lambda \dot{x},$$

where λ is a variable scalar and x and $\dot{x}$ are constant vectors since these are the vectors $x(t)$ and $\dot{x}(t)$ evaluated at the point P.

12.3 Osculating plane of a curve.

Given a curve C, let P, Q, R be three distinct points on C (Fig. 24). If the plane containing the points P, Q, R

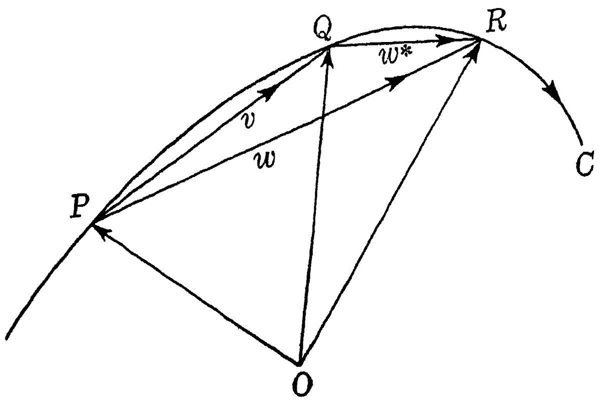

Fig. 24.

approaches a limiting position as Q and R approach P, this limiting plane is called the *osculating plane of the curve C at the point P.*

Let the equation of C be

$$x = x(t)$$

and let P, Q, R correspond to the values t, $t + \lambda$, $t + \mu$, respectively. Then the osculating plane is determined as the limiting position of the plane through P containing the vectors v and w, or v and w^*, where these vectors are defined as follows:

$$v = \frac{x(t + \lambda) - x(t)}{\lambda}$$

$$w = \frac{x(t + \mu) - x(t)}{\mu}$$

$$w^* = \frac{2(w - v)}{\mu - \lambda}.$$

We recall that if a scalar function $f(t)$ has suitable properties, it can be expanded as a Taylor Series with a remainder term; thus

$$f(a + t) = f(a) + f'(a)t + \frac{f''(a)}{1 \cdot 2}t^2 + R,$$

where the remainder term R is such that

$$\lim_{t \to 0} \frac{R}{t^2} = 0.$$

We assume that the vector function $x(t)$ is so expansible. Then

$$x(t + \lambda) = x(t) + \dot{x}(t)\lambda + \frac{\ddot{x}(t)}{1 \cdot 2}\lambda^2 + r_1,$$

where

$$\lim_{\lambda \to 0} \frac{r_1}{\lambda^2} = 0.$$

Now

$$v = \frac{x(t + \lambda) - x(t)}{\lambda} = \dot{x}(t) + \frac{\lambda}{2}\ddot{x}(t) + \frac{r_1}{\lambda}$$

and

$$w = \frac{x(t + \mu) - x(t)}{\mu} = \dot{x}(t) + \frac{\mu}{2}\ddot{x}(t) + \frac{r_2}{\mu}.$$

Hence

$$w^* = \frac{2(w - v)}{\mu - \lambda} = \ddot{x}(t) + r_3,$$

where

$$r_3 = \frac{2\left(\dfrac{r_2}{\mu} - \dfrac{r_1}{\lambda}\right)}{\mu - \lambda}.$$

For any position of the points Q, R we may write $\mu = \alpha\lambda$, where $\alpha \neq 0$; also $\alpha \neq 1$, since the points Q, R are distinct. We restrict α to the range $0 < \alpha_1 \leqq \alpha \leqq \alpha_2 < 1$. This means that λ, μ, and $\mu - \lambda$ are infinitesimals of the same order. By an infinitesimal is meant a variable which approaches zero as a limit. Two infinitesimals are said to be of the same order if their quotient approaches a (finite) number different from zero as a limit. Here

$$\frac{\mu - \lambda}{\mu} = 1 - \frac{\lambda}{\mu} = 1 - \frac{1}{\alpha},$$

and by the conditions imposed on α it follows that λ, μ, and $\mu - \lambda$ are infinitesimals of the same order. Now, since the limits of r_1/λ^2 and r_2/μ^2 are each zero, it follows that the limit of r_3 is zero.

Thus, in the limit

$$v = \dot{x}(t) \text{ and } w^* = \ddot{x}(t),$$

and the osculating plane is determined by these vectors provided they are linearly independent, which we assume to be the case. The position vector y of any point in the osculating plane is given by

$$y = x + \sigma_1\dot{x} + \sigma_2\ddot{x},$$

where σ_1, σ_2 are variable scalars and x, $\dot{x}$, $\ddot{x}$ are evaluated at the point P at which the osculating plane is taken.

Exercises

12.1. Show that the curve C defined by

$$\overrightarrow{OP} = x(t) = x_1(t)e_1 + x_2(t)e_2 + x_3(t)e_3$$

is a straight line if $x_1(t)$, $x_2(t)$, $x_3(t)$ are linear functions of t.

12.2. The curve which has as its equation

$$x = \alpha \cos t\, i + \alpha \sin t\, j + \beta t k,$$

with α, β positive constants, is called a *right circular helix.* Obtain the equation of the tangent line at an arbitrary point on the curve and show that the curve meets at a constant angle the generators of the cylinder on which it lies.

12.3. The curve whose equation is

$$x = t e_1 + t^2 e_2 + t^3 e_3$$

is called a *twisted cubic.* Obtain the equation of the tangent line at the point (2, 4, 8).

12.4. Show that the equation of the osculating plane can be written in the form of the box product

$$[y - x, \dot{x}, \ddot{x}] = 0.$$

12.5. Obtain the equation of the osculating plane of the right circular helix at an arbitrary point on it.

12.6. Obtain the equation of the osculating plane of the twisted cubic at the point (2, 4, 8).

12.7. A necessary and sufficient condition that a curve be a straight line is that $\dot{x}$ and $\ddot{x}$ be linearly dependent at an arbitrary point of the curve.

12.8. Show that the tangent and osculating plane of a curve at a point are invariant with respect to transformations on the parameter t, $t = \varphi(\tau)$, where $dt/d\tau \neq 0$.

12.4 Arc length of a curve.

References: *Blaschke* (26), p. 12; *Goursat* (34), p. 161.

The distance between two near-by points on a curve corresponding to t and $t + \Delta t$ measured along the straight line (chord) is given by

$$\begin{aligned} \Delta C &= \sqrt{\Delta x \cdot \Delta x} \\ &= \sqrt{\frac{\Delta x}{\Delta t} \cdot \frac{\Delta x}{\Delta t}}\Delta t. \end{aligned}$$

The *arc length* s of a curve joining two points A, B is defined to be the limit of the sum of the lengths of the sides of an

inscribed polygon as each side tends to zero. That is,

$$s = \lim_{N\to\infty} \sum_{i=1}^{N} \sqrt{\frac{\Delta \mathbf{x}_i}{\Delta t_i} \cdot \frac{\Delta \mathbf{x}_i}{\Delta t_i}}\,\Delta t_i \text{ as each } \Delta t_i \text{ approaches zero.}$$

But the limit of this sum is the definite integral

$$\int_{t_1}^{t_2} \sqrt{\dot{\mathbf{x}} \cdot \dot{\mathbf{x}}}\, dt,$$

where $\mathbf{x}(t_1) = \overrightarrow{OA}$ and $\mathbf{x}(t_2) = \overrightarrow{OB}$.

The arc length from a fixed point A to a variable point P along a curve, where $\mathbf{x}(t) = \overrightarrow{OP}$, is then given by

$$s(t) = \int_{t_1}^{t} \sqrt{\dot{\mathbf{x}} \cdot \dot{\mathbf{x}}}\, dt.$$

By the Fundamental Theorem of the integral calculus,

$$\frac{ds}{dt} = \sqrt{\dot{\mathbf{x}} \cdot \dot{\mathbf{x}}}.$$

If[1] $\dot{\mathbf{x}} \cdot \dot{\mathbf{x}} \overset{t}{\equiv} 1$, then $s = t - t_0$ and t is the arc length of the curve measured from a fixed point. Conversely, if t is arc length, then $ds/dt \overset{t}{\equiv} 1$, and $\dot{\mathbf{x}} \cdot \dot{\mathbf{x}} \overset{t}{\equiv} 1$. Hence *a necessary and sufficient condition that* $\mathbf{x}(t)$ *shall be a curve referred to its arc length as independent variable (or parameter) is that* $d\mathbf{x}/dt$ *shall be a unit vector at an arbitrary point of the curve.* We shall denote arc length as parameter by s and indicate derivatives with respect to s by primes, $\mathbf{x}' = d\mathbf{x}/ds$, $\mathbf{x}'' = d^2\mathbf{x}/ds^2, \cdots$.

From a practical standpoint two serious difficulties may be encountered in actually obtaining the equation of a curve in terms of its arc length as parameter. In the first place it may be impracticable to perform the indicated integration. If this is possible, it gives s explicitly in terms of t, say, $s = \varphi(t)$. The second difficulty is likely to be that of solving this equation for t in terms of s. However, exist-

[1] The notation $\overset{t}{\equiv}$ indicates an equality holding identically in t; that is, for every choice of t in the interval under consideration.

ence theorems (cf. *Goursat* (34), p. 35) assure us that such a transformation from t to s exists. For theoretical considerations the use of arc length as parameter often leads to greatly simplified expressions in connection with space curves.

Exercises

12.9. Obtain the arc length of the right circular helix from $t = 0$ to t, and deduce the equation of the curve in terms of arc length as parameter.

12.10. If the twisted cubic is given in terms of an $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ system,

$$\boldsymbol{x} = t\boldsymbol{i} + t^2\boldsymbol{j} + t^3\boldsymbol{k},$$

what difficulty do you meet in attempting to express the equation of the curve in terms of its arc length?

12.11. Consider the same problem as **12.10** for the curve whose equation is

$$\boldsymbol{x} = e^t\boldsymbol{i} + 2\sqrt{6}(t - 2)e^{t/2}\boldsymbol{j} + t^3\boldsymbol{k},$$

e being the base of natural logarithms.

12.12. If s is arc length of a curve, a necessary and sufficient condition that s^* also serve as arc length is that

$$s^* = \pm s + \lambda,$$

where λ is an arbitrary constant.

12.5 Curvature of a curve.

Descriptively speaking, the curvature of a curve at a point is a measure of the rate of departure (with respect to arc length) of the curve from its tangent line at that point. Let $\Delta\theta$ be the angle (in radian measure) between two near-by tangents corresponding to s and $s + \Delta s$, where the curve is expressed in terms of its arc length s:

$$C: \quad \boldsymbol{x} = \boldsymbol{x}(s).$$

The *curvature* $1/\rho$ of C at the point P, where P corresponds to the value s, is defined by

$$\frac{1}{\rho} = \lim_{\Delta s \to 0} \frac{\Delta\theta}{\Delta s} = \frac{d\theta}{ds}.$$

We observe that the curvature is a scalar quantity. To arrive at an analytical expression for it, it is convenient to introduce the spherical indicatrix of the tangents. Let the unit tangent vectors of the curve C be laid off from a fixed point. The end points of these vectors then form a curve which lies on a sphere of radius 1, and which is called the *spherical indicatrix of the tangents* of C. Suppose C is not a straight line, and consider a neighborhood of the curve at P such that no two tangents of C are parallel. Then the spherical indicatrix of this portion of C will be a non-closed curve whose points are in a one-to-one reciprocal correspondence with the points of C. Let P and Q on C at arc distance Δs apart have as correspondents P_1 and Q_1, respectively, on the spherical indicatrix. Let Δs_1 denote the length of arc of the spherical indicatrix from P_1 to Q_1. Let ΔC_1 be the arc length of the great circle between P_1 and Q_1. We assume that

$$\lim_{\Delta s_1 \to 0} \frac{\Delta C_1}{\Delta s_1} = 1.$$

But

$$\Delta C_1 = \Delta\theta.$$

Hence

$$\frac{1}{\rho} = \lim_{\Delta s \to 0} \frac{\Delta\theta}{\Delta s} = \lim_{\Delta s \to 0} \frac{\Delta C_1}{\Delta s} = \lim_{\Delta s \to 0} \frac{\Delta C_1}{\Delta s_1} \frac{\Delta s_1}{\Delta s};$$

that is,

$$\frac{1}{\rho} = \frac{ds_1}{ds}.$$

But the *position vector of points* on the spherical indicatrix is simply $\mathbf{x}'$. Hence

$$(ds_1)^2 = d\mathbf{x}' \cdot d\mathbf{x}',$$

and therefore

$$\frac{1}{\rho} = \sqrt{\frac{d\mathbf{x}'}{ds} \cdot \frac{d\mathbf{x}'}{ds}} = \sqrt{\mathbf{x}'' \cdot \mathbf{x}''}.$$

The vector $\mathbf{x}''$ is called the *curvature vector;* its length is the curvature of the curve. Since

$$\mathbf{x}' \cdot \mathbf{x}' \overset{s}{\equiv} 1,$$

we have by differentiation

$$x' \cdot x'' \overset{s}{\equiv} 0.$$

That is, the *curvature vector is a vector which lies in the osculating plane and which is perpendicular to the tangent vector.*

12.6 Principal normal and torsion.

A line passing through a point of a curve C, and perpendicular to the tangent at that point, is called a *normal* to the curve. That normal to a curve which lies in the osculating plane is called the *principal normal.* The normal which is perpendicular to the osculating plane is called the *binormal* of the curve.

We now set up at each point of a curve a system of three unitary orthogonal vectors ξ_1, ξ_2, ξ_3, defined as follows:

$$\begin{aligned} \xi_1 &= \text{unit tangent vector} \\ \xi_2 &= \text{unit principal normal vector} \\ \xi_3 = \xi_1 \times \xi_2 &= \text{unit binormal vector.} \end{aligned}$$

Descriptively speaking, the torsion of a curve is a measure of the rate of departure of a curve from its osculating plane. Let $\Delta\varphi$ denote the angle between two near-by binormals at points on the curve corresponding to s and $s + \Delta s$. The *torsion* $1/\tau$ is defined by

$$\frac{1}{\tau} = \lim_{\Delta s \to 0} \frac{\Delta\varphi}{\Delta s}.$$

Introducing the spherical indicatrix of the binormals, one obtains

$$\frac{1}{\tau} = \sqrt{\frac{d\xi_3}{ds} \cdot \frac{d\xi_3}{ds}} = \sqrt{\xi_3' \cdot \xi_3'}.$$

12.7 The Frenet formulas.

The set of vectors ξ_1, ξ_2, ξ_3 constitute a basis for the vector space. Since $\xi_1 = x'$ and $\xi_2 = \rho x''$, it follows that

$$\xi_1' = \frac{\xi_2}{\rho}.$$

We now similarly express the vectors ξ_2' and ξ_3' in terms of the vectors ξ_1, ξ_2, ξ_3. Let, then,

$$\xi_3' = \alpha\xi_1 + \beta\xi_2 + \gamma\xi_3,$$

where α, β, γ are now to be determined. Since

$$\xi_3 \cdot \xi_3 \stackrel{s}{\equiv} 1,$$

we have by differentiation $\xi_3 \cdot \xi_3' \stackrel{s}{\equiv} 0$. Operate on the above equation by $\xi_3 \cdot \xi_3$. Then

$$0 = \xi_3 \cdot \xi_3' = \alpha \cdot 0 + \beta \cdot 0 + \gamma \cdot 1,$$

whence $\gamma = 0$. From the identity $\xi_1 \cdot \xi_3 \stackrel{s}{\equiv} 0$ we have by differentiation

$$\xi_1' \cdot \xi_3 + \xi_1 \cdot \xi_3' \equiv 0$$

or

$$\frac{\xi_2}{\rho} \cdot \xi_3 + \xi_1 \cdot \xi_3' \equiv 0; \text{ that is, } \xi_1 \cdot \xi_3' \equiv 0.$$

Hence in the above $\alpha = 0$, and we have

$$\xi_3' = \beta\xi_2.$$

Forming the scalar product of each side of this equation gives

$$\xi_3' \cdot \xi_3' = \beta^2\xi_2 \cdot \xi_2;$$

that is, $1/\tau^2 = \beta^2$, whence $\beta = \pm 1/\tau$. We select the sign so that

$$\xi_3' = -\frac{\xi_2}{\tau}.$$

From the relation $\xi_2 = \xi_3 \times \xi_1$ one can now obtain ξ_2'.

Thus we have the relations

$$\begin{aligned} \xi_1' &= \frac{\xi_2}{\rho} \\ \xi_2' &= -\frac{\xi_1}{\rho} + \frac{\xi_3}{\tau} \\ \xi_3' &= -\frac{\xi_2}{\tau}, \end{aligned}$$

known as the **Frenet formulas** (1847). These formulas are

fundamental in the differential geometry theory of space curves.

Exercises

12.13. A necessary and sufficient condition that a curve be a straight line is that its curvature be zero at every point.

12.14. A necessary and sufficient condition that a curve, not straight, be a plane curve is that its torsion vanish identically.

12.15. Verify the Frenet formulas for the right circular helix. Observe that the curvature and torsion are each constant for this curve.

12.16. Investigate the spherical indicatrices of the tangent, normal, and binormal of the right circular helix.

12.17. By means of the Frenet formulas, or otherwise, show that

$$\frac{1}{\tau} = \frac{[\boldsymbol{x}'\boldsymbol{x}''\boldsymbol{x}''']}{\boldsymbol{x}'' \cdot \boldsymbol{x}''}.$$

12.18. In terms of a general parameter t, show that

$$\frac{1}{\rho^2} = \frac{(\dot{\boldsymbol{x}} \times \ddot{\boldsymbol{x}}) \cdot (\dot{\boldsymbol{x}} \times \ddot{\boldsymbol{x}})}{(\dot{\boldsymbol{x}} \cdot \dot{\boldsymbol{x}})^3}$$

$$\frac{1}{\tau} = \frac{[\dot{\boldsymbol{x}}\ddot{\boldsymbol{x}}\dddot{\boldsymbol{x}}]}{(\dot{\boldsymbol{x}} \times \ddot{\boldsymbol{x}}) \cdot (\dot{\boldsymbol{x}} \times \ddot{\boldsymbol{x}})}.$$

12.19. If all the osculating planes of a curve pass through a fixed point, the curve is plane.

12.20. If $\boldsymbol{x}$ and $\boldsymbol{x} + \Delta\boldsymbol{x}$ are position vectors of two neighboring points corresponding to s and $s + \Delta s$ on a *space* curve; that is, $1/\rho \neq 0$, $1/\tau \neq 0$, show that

(1) $\boldsymbol{\xi}_1 \cdot \Delta\boldsymbol{x}$ is an infinitesimal of the same order as Δs.

(2) $\boldsymbol{\xi}_2 \cdot \Delta\boldsymbol{x}$ is an infinitesimal of the second order with respect to Δs.

(3) $\boldsymbol{\xi}_3 \cdot \Delta\boldsymbol{x}$ is an infinitesimal of the third order with respect to Δs.

12.21. If the equation of a curve is

$$\boldsymbol{x} = x_1(t)\boldsymbol{e}_1 + x_2(t)\boldsymbol{e}_2 + x_3(t)\boldsymbol{e}_3,$$

where $x_1(t)$, $x_2(t)$, and $x_3(t)$ are quadratic functions of t, the curve is plane.

§13. Motion of a Particle

13.1 Velocity and acceleration vectors.

References: *Blaschke* (26), pp. 28–30; *Webster* (52) pp. 9–17.

A curve may be regarded as the path of a moving particle whose position is known as a function of the time. The arc length s of the particle measured from some initial point is likewise a function of the time t. We take t as the independent variable. Then

$$\frac{d\mathbf{x}}{dt} = \frac{d\mathbf{x}}{ds}\frac{ds}{dt} = \frac{ds}{dt}\boldsymbol{\xi}_1,$$

that is, the *velocity vector* $d\mathbf{x}/dt$ has the same direction as the tangent vector; its length is ds/dt. The scalar ds/dt is called the *speed* of the particle.

The vector which gives the rate of change of the velocity vector with respect to the time is called the *acceleration vector*. It is given by

$$\frac{d^2\mathbf{x}}{dt^2} = \frac{d^2\mathbf{x}}{ds^2}\left(\frac{ds}{dt}\right)^2 + \frac{d\mathbf{x}}{ds}\frac{d^2s}{dt^2},$$

which is a vector in the osculating plane of the curve traversed by the moving particle. By the Frenet formulas,

$$\frac{d^2\mathbf{x}}{dt^2} = \frac{\left(\dfrac{ds}{dt}\right)^2}{\rho}\boldsymbol{\xi}_2 + \frac{d^2s}{dt^2}\boldsymbol{\xi}_1,$$

which is the resolution of the acceleration vector along the tangent and principal normal of the curve. The coefficient of $\boldsymbol{\xi}_1$, viz., d^2s/dt^2 is called the *tangential acceleration* of the particle. The velocity is wholly a tangential velocity, whereas part of the acceleration is used to effect the curvature of the path. We observe that the normal component of the acceleration is zero if and only if the path is a straight line. If the motion of a particle along any path has a uniform speed; that is, if ds/dt = constant, then the tangential acceleration is zero, and conversely.

13.2 Axis of rotation of a rigid body.

If a rigid body rotating about an axis turns through an angle $\Delta\varphi$ in the time Δt, the limit of $\Delta\varphi/\Delta t$ as Δt approaches zero, viz., $d\varphi/dt$, is called the *angular speed of rotation*. The *angular velocity vector* can be introduced just as in **§10.5.**[2] If P is an arbitrary point of the rigid body whose position vector with respect to a point on the axis of rotation is $\boldsymbol{x}$, we have

$$\frac{d\boldsymbol{x}}{dt} = \boldsymbol{r} \times \boldsymbol{x},$$

where $\boldsymbol{r}$ is the angular velocity vector.

The system of vectors $\boldsymbol{\xi}_1$, $\boldsymbol{\xi}_2$, $\boldsymbol{\xi}_3$ defined in **§12.6** which are associated with each point of a curve is called the *local trihedral*. Since the lengths of these vectors and the angles between them remain constant, the local trihedral may be regarded as a moving rigid body. Let us determine its angular velocity vector. In order to consider the *rotation* properties of the local trihedral, let it be set up at a *fixed* point. Since the relation $d\boldsymbol{x}/dt = \boldsymbol{r} \times \boldsymbol{x}$ holds for an arbitrary point P of the rigid body, it holds in particular for the end points of the vectors $\boldsymbol{\xi}_1$, $\boldsymbol{\xi}_2$, $\boldsymbol{\xi}_3$. Hence

$$\frac{d\boldsymbol{\xi}_1}{dt} = \boldsymbol{r} \times \boldsymbol{\xi}_1, \qquad \frac{d\boldsymbol{\xi}_2}{dt} = \boldsymbol{r} \times \boldsymbol{\xi}_2, \qquad \frac{d\boldsymbol{\xi}_3}{dt} = \boldsymbol{r} \times \boldsymbol{\xi}_3.$$

Let $\boldsymbol{r}$ be expressed in the form

$$\boldsymbol{r} = \alpha\boldsymbol{\xi}_1 + \beta\boldsymbol{\xi}_2 + \gamma\boldsymbol{\xi}_3.$$

By means of the Frenet formulas,

$$\frac{d\boldsymbol{\xi}_1}{dt} = \frac{d\boldsymbol{\xi}_1}{ds}\frac{ds}{dt} = \frac{ds}{dt}\frac{\boldsymbol{\xi}_2}{\rho} = -\beta\boldsymbol{\xi}_3 + \gamma\boldsymbol{\xi}_2.$$

Hence

$$\beta = 0, \gamma = \frac{ds}{dt}\frac{1}{\rho}.$$

[2] A detailed discussion of the angular velocity vector is given later; see p. 84.

Similarly one deals with the other equations. The result is

$$r = \frac{ds}{dt}\left(\frac{1}{\tau}\xi_1 + \frac{1}{\rho}\xi_3\right),$$

a vector in the plane of the tangent and binormal. If the torsion is zero, the rotation is about the binormal, the rate of rotation being directly proportional to the curvature and the speed at which the trihedral is being propagated along the curve. This formula shows clearly that the curvature is a measure of the rate of rotation of the trihedral about the binormal and that the torsion measures its rate of rotation about the tangent line.

Exercises

13.1. If $\boldsymbol{x} \cdot d\boldsymbol{x} = 0$, show that $|\boldsymbol{x}|$ is constant. If $\boldsymbol{x} \times d\boldsymbol{x} = \mathbf{0}$, show that $\boldsymbol{x}$ remains parallel to itself. If $\boldsymbol{x} \cdot (d\boldsymbol{x} \times d^2\boldsymbol{x}) = 0$, show that $\boldsymbol{x} \times d\boldsymbol{x}$ has a fixed direction, and that $\boldsymbol{x}$ is parallel to a fixed plane. (Assume $\boldsymbol{x}$ depends upon a single scalar variable.)

13.2. Show that the orbit of a particle subject only to a central force is a plane curve. Deduce Kepler's first law of motion; viz., that the radius vector of the particle, with respect to the position of the central force as origin, sweeps over equal areas in equal times.

13.3. The product of the mass of a particle and its velocity vector is called its *linear momentum vector;* the moment of the linear momentum vector with respect to a point is called the *angular momentum vector*. If a particle is acted on only by a central force, show that its angular momentum vector with respect to the origin of the central force is a constant vector.

13.4. If a particle is subject only to a central force, the magnitude of its velocity vector is the same at all points equally distant from the origin of the central force; that is, the speed of the particle is a function only of $\boldsymbol{x} \cdot \boldsymbol{x}$ and the initial conditions.

13.3 Moving coördinate system.

References: *Wills* (23), pp. 53–56; *Ames-Murnaghan* (24), pp. 88–102; *Eisenhart* (32), p. 30 *ff.*; *Haas* (9), pp. 23–35.

Thus far in our study of a moving point, or of variable vectors, we have supposed the objects referred to a *fixed* coördinate system determined by a point O and a system of *constant* (or fixed) base vectors. However, in the analysis of the motion of a particle the employment of a moving coördinate system is frequently a powerful device.

Let the point $\bar{O}$ and constant base vectors $\bar{\mathbf{e}}_1$, $\bar{\mathbf{e}}_2$, $\bar{\mathbf{e}}_3$ determine a fixed coördinate system, and let the coördinates of a point with respect to this system be denoted by (y_1, y_2, y_3). Let there also be given at each point of the space under consideration a "local" system of base vectors, say $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$. Since these vectors may vary from point to point, they must be regarded as functions of the coördinates y_1, y_2, y_3. Thus

$$\mathbf{e}_1 = \alpha(y_1, y_2, y_3)\bar{\mathbf{e}}_1 + \beta(y_1, y_2, y_3)\bar{\mathbf{e}}_2 + \gamma(y_1, y_2, y_3)\bar{\mathbf{e}}_3,$$

and similarly for the vectors $\mathbf{e}_2$ and $\mathbf{e}_3$. Let the origin O of

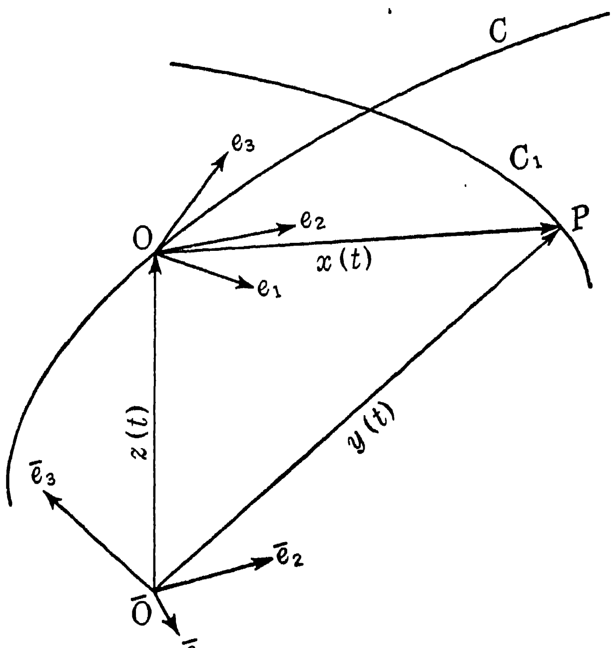

Fig. 25.

the moving coördinate system determined by O; $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$ traverse a curve C whose equation with respect to the fixed coördinate system is

$$\overrightarrow{\bar{O}O} = \boldsymbol{z} = y_1(t)\bar{\boldsymbol{e}}_1 + y_2(t)\bar{\boldsymbol{e}}_2 + y_3(t)\bar{\boldsymbol{e}}_3.$$

Then along the curve C the coördinates of a point are known as functions of a single vaarible t. Hence the local system of base vectors at points of C are known as functions of t, which are given by substituting in the coefficients α, β, etc., the y_1, y_2, and y_3 in terms of t.

Consider now a moving point P whose position vector with respect to the fixed system is $\boldsymbol{y}(t)$, and with respect to the moving system is $\boldsymbol{x}(t)$. If $\boldsymbol{z}(t)$ is the position vector of O with respect to the fixed system, we have

$$\boldsymbol{y}(t) = \boldsymbol{z}(t) + \boldsymbol{x}(t).$$

Differentiating with respect to t, we have

$$\frac{d\boldsymbol{y}(t)}{dt} = \frac{d\boldsymbol{z}(t)}{dt} + \frac{d\boldsymbol{x}(t)}{dt}.$$

In this we keep in mind that the vectors $\boldsymbol{y}(t)$ and $\boldsymbol{z}(t)$ are variable only because of their variable coefficients, but that $\boldsymbol{x}(t)$ varies not only because of its variable coefficients but also because it is expressed in terms of variable base vectors. In order to exhibit the situation more clearly, we write

$$\boldsymbol{y}(t) = \boldsymbol{z}(t) + \{\alpha_1(t)\boldsymbol{e}_1(t) + \alpha_2(t)\boldsymbol{e}_2(t) + \alpha_3(t)\boldsymbol{e}_3(t)\}.$$

Then, applying the rule for differentiating a product of a scalar times a vector, we have

$$\frac{d\boldsymbol{y}}{dt} = \frac{d\boldsymbol{z}}{dt} + \left(\frac{d\alpha_1}{dt}\boldsymbol{e}_1 + \frac{d\alpha_2}{dt}\boldsymbol{e}_2 + \frac{d\alpha_3}{dt}\boldsymbol{e}_3\right) + \left(\alpha_1\frac{d\boldsymbol{e}_1}{dt} + \alpha_2\frac{d\boldsymbol{e}_2}{dt} + \alpha_3\frac{d\boldsymbol{e}_3}{dt}\right).$$

Let the vectors $d\boldsymbol{e}_1/dt$, $d\boldsymbol{e}_2/dt$, $d\boldsymbol{e}_3/dt$ be expressed in terms of the local base vectors $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$. Then we would have a

relation of the form

$$\frac{d\boldsymbol{y}}{dt} = \frac{d\boldsymbol{z}}{dt} + \{\beta_1(t)\boldsymbol{e}_1(t) + \beta_2(t)\boldsymbol{e}_2(t) + \beta_3(t)\boldsymbol{e}_3(t)\},$$

where the β_1, β_2, β_3 are known functions of t. One could now differentiate again with respect to t and eliminate the derivatives $d\boldsymbol{e}_1/dt$, $d\boldsymbol{e}_2/dt$, $d\boldsymbol{e}_3/dt$ as before. We see that repeated differentiations with respect to t introduce no logical difficulties.

The moving point P traverses a curve C_1 and the variable t establishes a correspondence between the points of C and C_1, O and P being corresponding points. The equation

$$\frac{d\boldsymbol{y}}{dt} = \frac{d\boldsymbol{z}}{dt} + (\beta_1\boldsymbol{e}_1 + \beta_2\boldsymbol{e}_2 + \beta_3\boldsymbol{e}_3)$$

relates the tangent vectors to C and C_1 at corresponding points; or, if t is interpreted as the time, it relates the velocity vectors of O and P.

As an important special case, let C be referred to its arc length s, and let $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$ be the system of vectors $\boldsymbol{\xi}_1$, $\boldsymbol{\xi}_2$, $\boldsymbol{\xi}_3$ defined in **§12.6**;

$$\boldsymbol{y}(s) = \boldsymbol{z}(s) + \alpha_1(s)\boldsymbol{\xi}_1(s) + \alpha_2(s)\boldsymbol{\xi}_2(s) + \alpha_3(s)\boldsymbol{\xi}_3(s).$$

Then

$$\frac{d\boldsymbol{y}}{ds} = \frac{d\boldsymbol{z}}{ds} + \alpha_1'\boldsymbol{\xi}_1 + \alpha_2'\boldsymbol{\xi}_2 + \alpha_3'\boldsymbol{\xi}_3 + \alpha_1\boldsymbol{\xi}_1' + \alpha_2\boldsymbol{\xi}_2' + \alpha_3\boldsymbol{\xi}_3',$$

the primes denoting derivatives with respect to s. If we replace $d\boldsymbol{z}/ds$ by $\boldsymbol{\xi}_1$ and make use of the Frenet formulas, the equation reduces to

$$(A)\quad \frac{d\boldsymbol{y}}{ds} = \left(1 + \alpha_1' - \frac{\alpha_2}{\rho}\right)\boldsymbol{\xi}_1 + \left(\frac{\alpha_1}{\rho} + \alpha_2' - \frac{\alpha_3}{\tau}\right)\boldsymbol{\xi}_2 + \left(\alpha_3' + \frac{\alpha_2}{\tau}\right)\boldsymbol{\xi}_3.$$

We have seen (**§10.6**) that a finite rotation about a line cannot be represented by a vector which is parallel to the

line and whose length is the measure of the rotation. We now show, however, that such a representation is valid for an infinitesimal rotation, or an angular velocity.

S_1, S_2, S_3 are three rigid systems. S_3 rotates relatively to S_2 about the (directed) line a_{32} with angular velocity ω_{32}; S_2 relatively to S_1 about a_{21} with angular velocity ω_{21}. Moreover a_{32} and a_{21} meet in a point O. It is required to describe the motion of S_3 relative to S_1.

The point O clearly remains fixed relatively to S_1. Thus the resulting motion is a rotation about a line through O. The axis a_{31} will contain, besides O, any point such that

$$\boldsymbol{V}_{31} = \boldsymbol{V}_{32} + \boldsymbol{V}_{21} = \boldsymbol{0},$$

where $\boldsymbol{V}_{ij}$ is the linear velocity of the point (thought of as a point of S_i) relative to S_j.

Unless both ω_{32} and ω_{21} are zero, only points in the plane of a_{32} and a_{21} can have a zero velocity, $\boldsymbol{V}_{31} = \boldsymbol{0}$. If A_3 is such a point, different from O, then for it the vector equation for compounding linear velocities yields

$$v_{31} = v_{32} + v_{21} = 0,$$

where the v_{ij} are speeds. Let l_3, l_2 denote the distances from A_3 to a_{32} and a_{21}, respectively. Then

$$v_{31} = l_3\omega_{32} - l_2\omega_{21} = 0. \tag{1}$$

Let the plane containing the three axes of rotation be oriented and let $\alpha_1 \neq 0$ denote the angle measured from a_{21} to a_{32}. Let α_2 be the angle from a_{31} to a_{21}, and let α_3 be the angle from a_{32} to a_{31}. Then

$$\alpha_1 + \alpha_2 + \alpha_3 = 0,.$$

or

$$\alpha = \alpha_2 + \alpha_3, \tag{2}$$

where $\alpha = -\alpha_1$. Equation (1) may now be written

$$\omega_{32} \sin \alpha_3 = \omega_{21} \sin \alpha_2. \tag{3}$$

From (2) and (3) we obtain

$$(4) \qquad \begin{cases} \tan \alpha_2 = \dfrac{\omega_{32} \sin \alpha}{\omega_{32} \cos \alpha + \omega_{21}} \\ \tan \alpha_3 = \dfrac{\omega_{21} \sin \alpha}{\omega_{32} + \omega_{21} \cos \alpha}. \end{cases}$$

Now each point B_2 of a_{32} (thought of as a point of S_2) has, relative to S_1, just the same velocity as the point B_3 of S_3 which momentarily coincides with it. Let $\overline{B_2E}$, $\overline{B_3C}$ be the perpendiculars from $B_2 = B_3$ to a_{21} and a_{31}, respectively. Then for B_2, B_3

$$v_{21} = \overline{B_2E} \cdot \omega_{21}$$
$$v_{31} = \overline{B_3C} \cdot \omega_{31},$$

or

$$\omega_{31} = \omega_{21} \frac{\overline{B_2E}}{\overline{B_3C}} = \omega_{21} \frac{\sin \alpha}{\sin \alpha_3} = \omega_{32} \frac{\sin \alpha}{\sin \alpha_2}.$$

Replacing α_2, α_3 by means of (4),

$$\omega_{31} = \sqrt{\omega_{21}^2 + \omega_{32}^2 + 2\omega_{21}\omega_{32} \cos \alpha}.$$

This relation shows that our representation (§**10.5**) of angular velocities by directed line segments obeys the parallelogram law in compounding. The other requirements (§**3.3**) for a vector are evidently satisfied by the representation. Hence we conclude that an angular velocity, or an infinitesimal rotation, is a vector quantity.

We now consider a coördinate system rotating about a fixed point O. We suppose the vectors $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$ at O behave as a rigid body under the rotation. Let the independent variable t be interpreted as the time. For this case we have

$$y(t) = \mathbf{a} + x(t) = \mathbf{a} + \alpha_1(t)\boldsymbol{e}_1 + \alpha_2(t)\boldsymbol{e}_2 + \alpha_3(t)\boldsymbol{e}_3,$$

where $\boldsymbol{a}$ is the position vector of the fixed point O with respect to $\overline{O}$, the origin of the fixed coördinate system.

Differentiating with respect to t, one obtains

$$\frac{d\boldsymbol{y}}{dt} = \left(\frac{d\alpha_1}{dt}\boldsymbol{e}_1 + \frac{d\alpha_2}{dt}\boldsymbol{e}_2 + \frac{d\alpha_3}{dt}\boldsymbol{e}_3\right) + \left(\alpha_1\frac{d\boldsymbol{e}_1}{dt} + \alpha_2\frac{d\boldsymbol{e}_2}{dt} + \alpha_3\frac{d\boldsymbol{e}_3}{dt}\right).$$

The left member of this equation is the velocity vector of the point P as observed from the fixed reference system at O; the first term on the right is the velocity vector of P as determined by an observer in the rotating system; the second term on the right is the velocity vector which would be imparted to the particle P, regarded as fixed in the moving system, because of the rotation of the system. We have seen that this second term can be expressed in terms of the angular velocity vector, which we denote by $\boldsymbol{r}$. Then

$$\frac{d\boldsymbol{y}}{dt} = \left(\frac{d\alpha_1}{dt}\boldsymbol{e}_1 + \frac{d\alpha_2}{dt}\boldsymbol{e}_2 + \frac{d\alpha_3}{dt}\boldsymbol{e}_3\right) + (\boldsymbol{r} \times \boldsymbol{x}).$$

Differentiating again with respect to t,

$$\frac{d^2\boldsymbol{y}}{dt^2} = \left(\frac{d^2\alpha_1}{dt^2}\boldsymbol{e}_1 + \frac{d^2\alpha_2}{dt^2}\boldsymbol{e}_2 + \frac{d^2\alpha_3}{dt^2}\boldsymbol{e}_3\right) + \left(\frac{d\alpha_1}{dt}\frac{d\boldsymbol{e}_1}{dt} + \frac{d\alpha_2}{dt}\frac{d\boldsymbol{e}_2}{dt} + \frac{d\alpha_3}{dt}\frac{d\boldsymbol{e}_3}{dt}\right) + \left(\frac{d\boldsymbol{r}}{dt} \times \boldsymbol{x}\right) + \left(\boldsymbol{r} \times \frac{d\boldsymbol{x}}{dt}\right).$$

However, through the introduction of the angular velocity vector $\boldsymbol{r}$, any term of the form

$$\lambda_1\frac{d\boldsymbol{e}_1}{dt} + \lambda_2\frac{d\boldsymbol{e}_2}{dt} + \lambda_3\frac{d\boldsymbol{e}_3}{dt}$$

can be expressed as

$$\boldsymbol{r} \times (\lambda_1\boldsymbol{e}_1 + \lambda_2\boldsymbol{e}_2 + \lambda_3\boldsymbol{e}_3).$$

By this means the above equation is reducible to

$$(B) \quad \frac{d^2\boldsymbol{y}}{dt^2} = \frac{d^2\alpha_1}{dt^2}\boldsymbol{e}_1 + \frac{d^2\alpha_2}{dt^2}\boldsymbol{e}_2 + \frac{d^2\alpha_3}{dt^2}\boldsymbol{e}_3 + 2\boldsymbol{r} \times \left(\frac{d\alpha_1}{dt}\boldsymbol{e}_1 + \frac{d\alpha_2}{dt}\boldsymbol{e}_2 + \frac{d\alpha_3}{dt}\boldsymbol{e}_3\right) + \left(\frac{dr_1}{dt}\boldsymbol{e}_1 + \frac{dr_2}{dt}\boldsymbol{e}_2 + \frac{dr_3}{dt}\boldsymbol{e}_3\right) \times \boldsymbol{x} + \boldsymbol{r} \times (\boldsymbol{r} \times \boldsymbol{x}),$$

where r_1, r_2, r_3 are the coefficients of $\boldsymbol{r}$ expressed in terms of the local base vectors $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$.

Exercises

13.4. Let C be a curve referred to arc length as parameter, and let C_1 be the curve traced out by a point P which is on the tangent to C at O and such that the distance $\overline{OP}$ remains constant.

(1) Show that the tangent to C_1 is parallel to the osculating plane of C at the corresponding point.

(2) Show that the arc length s_1 of C_1 is given by

$$s_1 = \int_{s_0}^{s}\left[1 + \left(\frac{\alpha}{\rho}\right)^2\right]^{\frac{1}{2}} ds,$$

where α is the distance $\overline{OP}$, and $1/\rho$ is the curvature of C at O.

13.5. If a point Q on the tangent to a curve C at a variable point P moves so that its locus C_1 is always perpendicular to the tangent, the curve C_1 is called an *involute* of C. Show that the distance $\overline{PQ}$ is equal to $\lambda + s$, where λ is a constant and s is the arc length of C measured from some point. This is sometimes called the "string property" of the involute. Describe how an involute can be generated mechanically.

13.6. Let P be a variable point of a curve C, and let C_1 be a curve traced out by a point Q in the normal plane of C at P such that the curve C_1 is perpendicular to this plane.

(1) Show that the tangents to C_1 and C at corresponding points are parallel.

(2) Show that the distance $\overline{PQ}$ is constant (the curves C_1 and C are called *parallel curves*).

(3) The curve C bears the same relation to C_1 that C_1 does to C.

13.7. If two curves have the same tangent at corresponding points, they coincide.

13.8. If two curves have the same binormal at corresponding points, they coincide, or each is a plane curve.

13.9. If $\boldsymbol{r}$ denotes the position vector of a point P moving in a plane with respect to an origin O taken in the plane, show that the velocity vector of the point is given by

$$\frac{d\mathbf{r}}{dt} = \frac{dr}{dt}\mathbf{e}_1 + r\frac{d\theta}{dt}\mathbf{e}_2,$$

where r, θ are polar coördinates of P, and $\mathbf{e}_1$, $\mathbf{e}_2$ are unit vectors at P in the directions of r increasing, and θ increasing, respectively. Also show that the acceleration vector of the point is given by

$$\frac{d^2\mathbf{r}}{dt^2} = \left\{\frac{d^2r}{dt^2} - r\left(\frac{d\theta}{dt}\right)^2\right\}\mathbf{e}_1 + \left\{r\frac{d^2\theta}{dt^2} + 2\frac{dr}{dt}\frac{d\theta}{dt}\right\}\mathbf{e}_2.$$

13.10. Verify that equation (A) may be written in the form

$$\frac{d\mathbf{y}}{ds} = \frac{d\mathbf{z}}{ds} + (\alpha_1'\mathbf{e}_1 + \alpha_2'\mathbf{e}_2 + \alpha_3'\mathbf{e}_3) + \mathbf{r} \times \mathbf{x},$$

where $\mathbf{r}$ is the angular velocity vector of the local trihedral obtained in **§13.2.**

13.11. With reference to equation (B), establish the following:
(1) The term

$$2\mathbf{r} \times \left(\frac{d\alpha_1}{dt}\mathbf{e}_1 + \frac{d\alpha_2}{dt}\mathbf{e}_2 + \frac{d\alpha_3}{dt}\mathbf{e}_3\right) \quad .$$

is called the *acceleration of Coriolis;* it is perpendicular to the angular velocity vector and to the relative velocity vector, that is, the velocity vector as observed from the rotating system.

(2) If the point P is fixed in the moving system and $\mathbf{r}$ is constant, the total acceleration is given by

$$\mathbf{r} \times (\mathbf{r} \times \mathbf{x}).$$

This is called the *centripetal acceleration;* it is a vector directed from P toward the axis of revolution and has a magnitude directly proportional to the distance from P to the axis and to the square of the angular speed of rotation.

13.12. Under the hypothesis leading to equation (B),
(1) Show that

$$\mathbf{e}_\alpha \cdot \frac{d\mathbf{e}_\beta}{dt} + \mathbf{e}_\beta \cdot \frac{d\mathbf{e}_\alpha}{dt} = 0, \ \alpha, \beta = 1, 2, 3.$$

(2) Obtain an expression for the angular velocity vector $\mathbf{r}$.

§14. On the Geometry of a Surface

References: *Blaschke* (26); *Juvet* (11), Chapter III.

14.1 Notion of a surface.

Let P be a point whose position vector with respect to a fixed system is $\mathbf{x}(u, v)$ where u, v are independent scalar variables. The points P then form a *surface* S whose vector equation is

$$S: \quad \mathbf{x} = \mathbf{x}(u, v).$$

If v is held constant while u varies, the point P traces a curve on the surface, which is called a *u-parametric curve.*

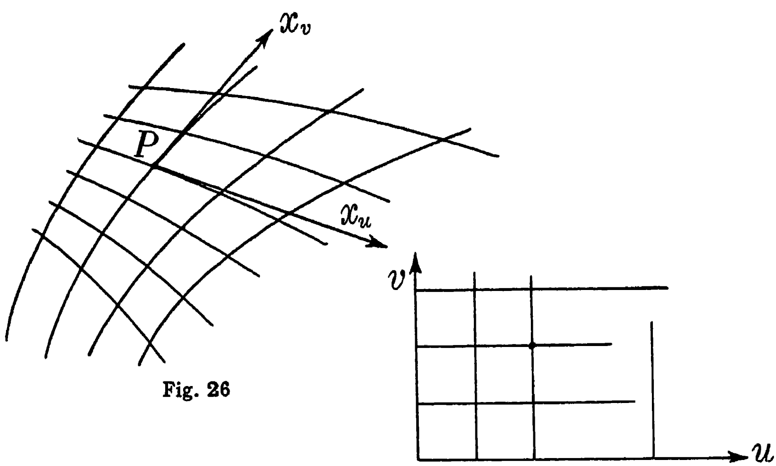

Fig. 26

Interchanging the role of u and v gives rise to another family of curves on the surface. Thus the surface is covered by a *parametric net;* through each point P of the surface there passes one curve of each family. The parametric net then constitutes a *coördinate system* on the surface, since when the values of u and v are given, the point P is uniquely determined, and conversely, as we shall assume.

The tangent vectors to the u- and v-parametric curves are given by $\mathbf{x}_u$ and $\mathbf{x}_v$, respectively, where the subscripts denote partial derivatives with respect to the variables

u and v. We *assume* $\mathbf{x}_u$ and $\mathbf{x}_v$ *to be linearly independent* at each point P of S under consideration.

Consider now an arbitrary curve C on S. Let it be determined by the equations

$$u = u(t), \; v = v(t).$$

Then

$$\frac{d\mathbf{x}}{dt} = \mathbf{x}_u \frac{du}{dt} + \mathbf{x}_v \frac{dv}{dt}$$

is the tangent vector to C. Since this vector is a linear combination of $\mathbf{x}_u$ and $\mathbf{x}_v$, it is in the plane of these vectors. This plane is called the *tangent plane* to S at the point P. The line at P which is orthogonal to the tangent plane at that point is called the *normal* to the surface. The unit normal vector $\boldsymbol{\xi}$ is then given by

$$\boldsymbol{\xi} = \frac{\mathbf{x}_u \times \mathbf{x}_v}{|\mathbf{x}_u \times \mathbf{x}_v|}.$$

14.2 First fundamental form.

Corresponding to an infinitesimal displacement in the three-space we have defined the differential of arc length ds by

$$(ds)^2 = d\mathbf{x} \cdot d\mathbf{x}.$$

Hence for a displacement $d\mathbf{x}$ on the surface

$$(ds)^2 = (\mathbf{x}_u du + \mathbf{x}_v dv) \cdot (\mathbf{x}_u du + \mathbf{x}_v dv)$$

or

$$(A) \qquad ds^2 = E\,du^2 + 2F\,du\,dv + G\,dv^2,$$

where we write ds^2 for $(ds)^2$ and E, F, G are defined by

$$E = \mathbf{x}_u \cdot \mathbf{x}_u, \; F = \mathbf{x}_u \cdot \mathbf{x}_v, \; G = \mathbf{x}_v \cdot \mathbf{x}_v.$$

The right member of (A) is known as the *first fundamental form* for the surface; it is a homogeneous, symmetric, quadratic differential form which is positive definite. The coefficients E, F, G of the form are of course generally functions of u and v.

We observe at once that a necessary and sufficient condition that the parametric net be orthogonal is that

$F \equiv 0$. The angle θ between the parametric curves is given by

$$\cos\theta = \frac{F}{\sqrt{E}\sqrt{G}}.$$

14.3 Element of area.

Consider the area bounded by the parametric curves corresponding to the values u, $u + du$, v, $v + dv$. If we neglect infinitesimals of higher order with respect to du and dv, this area is given by

$$d\sigma = ds_1 ds_2 \sin\theta,$$

where ds_1 and ds_2 are the arc lengths corresponding to the displacements specified by du and dv, respectively. But from the first fundamental form,

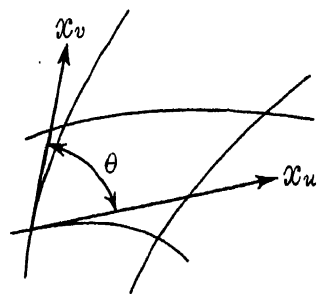

Fig. 27

$$ds_1 = \sqrt{E}\,du,\ ds_2 = \sqrt{G}\,dv.$$

From the expression for $\cos\theta$ we obtain

$$\sin\theta = \frac{\sqrt{EG - F^2}}{\sqrt{E}\sqrt{G}}.$$

Hence $d\sigma$, the *element of area*, is given by

$$d\sigma = \sqrt{EG - F^2}\,du\,dv.$$

14.4 Coördinate equation of a surface.

Suppose a surface S is defined by the equation

$$F(x, y, z) = 0,$$

where x, y, z are rectangular Cartesian coördinates in space. Let C be a curve defined by

$$x = x(t),\ y = y(t),\ z = z(t).$$

Then $\dot{x}$, $\dot{y}$, $\dot{z}$ are proportional to the direction cosines of the tangent to the curve, where the dots denote derivatives with respect to t. Suppose now the curve C is on the

surface S. This means that

$$F(x(t), y(t), z(t)) \stackrel{t}{\equiv} 0.$$

Differentiating this identity with respect to t, we obtain

$$\frac{\partial F}{\partial x}\dot{x} + \frac{\partial F}{\partial y}\dot{y} + \frac{\partial F}{\partial z}\dot{z} \stackrel{t}{\equiv} 0.$$

Therefore $\partial F/\partial x$, $\partial F/\partial y$, $\partial F/\partial z$ are proportional to the direction cosines of a direction which is perpendicular to the tangent to C. Since C may be *any* curve on S passing through a point P, it follows that $\partial F/\partial x$, $\partial F/\partial y$, $\partial F/\partial z$ are proportional to the direction cosines of the *normal* to S at the point P.

From the equation of the surface

$$F(x, y, z) = 0$$

we obtain by taking the total differential of each side

$$\frac{\partial F}{\partial x}dx + \frac{\partial F}{\partial y}dy + \frac{\partial F}{\partial z}dz = 0,$$

which equation must be satisfied for any differential displacement on the surface S specified by dx, dy, dz.

Exercises

14.1. The vector equation of a right circular cylinder about the k-axis is

$$\boldsymbol{x}(u, v) = \alpha \cos u\boldsymbol{i} + \alpha \sin u\boldsymbol{j} + \beta v\,\boldsymbol{k}.$$

Show that any helix on this cylinder has the property that the principal normal of the curve is also normal to the cylinder on which the curve lies. (Any curve possessing this property is a *geodesic* on the surface.)

14.2. The vector equation of a sphere of radius R whose center is at the origin is

$$\boldsymbol{x}(u, v) = R \cos u \cos v\boldsymbol{i} + R \cos u \sin v\boldsymbol{j} + R \sin u\,\boldsymbol{k},$$

where u is latitude and v is longitude. Obtain the first fundamental form for the surface. Obtain the element of area of the surface, and by integration find the area of the sphere.

14.3. The unit normal ξ to the surface is given by

$$\xi = \frac{\boldsymbol{x}_u \times \boldsymbol{x}_v}{\sqrt{EG - F^2}}.$$

14.4. From the fact that the first fundamental form (A) is positive definite, deduce that $EG - F^2$ is positive.

14.5. Show that the poles are singular points in the particular parametric representation of the sphere as given by Exercise **14.2**; that is, at these points $\boldsymbol{x}_u$ and $\boldsymbol{x}_v$ fail to be linearly independent.

§15. Scalar and Vector Fields

References: *Juvet* (11), pp. 77–83; *Kellogg* (40), Chapter II.

In the present section we suppose an $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ system, which together with a point O constitutes a rectangular Cartesian coördinate system. We shall denote the coördinates of a point by (x, y, z).

15.1 Notion of a field.

Let R be a region of space, and let there be assigned to each point P of R a real number. This number is a function of P which we denote by the symbol $f(P)$. The function $f(P)$ is called a *scalar point function* or a *scalar function of position*. The points of R together with the functional values $f(P)$ will be called a *scalar field over* R. Thus a given $f(P)$ defined at each point of R determines uniquely a scalar field over that region. Since there exists a one-to-one reciprocal correspondence between the points of R and ordered triples of numbers (x, y, z), where the ranges of x, y, and z are suitably restricted, $f(P)$ is equivalent to a function F of the variables x, y, and z:

$$f(P) \equiv F(x, y, z).$$

The temperature of points in the atmosphere at a given instant is evidently an example of a scalar field; the density of the atmosphere at a given time provides another scalar field over the same region.

A scalar field $f(P)$ defined over R is *continuous* at P_0 in R, if, given an arbitrary positive number ϵ, there exists a positive number η such that for all P in R satisfying the inequality $|\overrightarrow{P_0P}| < \eta$ it is true that $|f(P) - f(P_0)| < \epsilon$. The field is said to be continuous in R if it is continuous at each point of R. We shall suppose not only that the fields under discussion are continuous, except when otherwise noted, but that they possess such derivatives as may be needed.

Let $f(P)$ define a field over R, and let P_0 be a point of R. Then $f(P_0)$ is a real number, and all points P of R at which $f(P)$ has the same value are specified by

$$f(P) = f(P_0).$$

The set of points satisfying an equation of the form

$$f(P) = \text{constant}$$

constitutes a surface, which is called a *level surface* of the field.

If with each point P of a region R there is associated a vector $V(P)$, the points of R together with these vectors constitute a vector field over R. An example of a vector field is furnished by the wind velocity of points in the atmosphere at a given time. The definition of continuity of a vector field is similar to that of a scalar field.

Exercises

15.1. Give additional examples of a scalar field, and of a vector field. Consider the customary names of the level surfaces for the scalar fields which you have selected.

15.2. If $V(P)$ is given by

$$V(P) = \alpha(x, y, z)\boldsymbol{i} + \beta(x, y, z)\boldsymbol{j} + \gamma(x, y, z)\boldsymbol{k},$$

a necessary and sufficient condition that $V(P)$ be a continuous vector field over a region R is that the scalar coefficients be continuous scalar fields over R.

15.3. If $V(P)$ is a vector field over R, then $|V(P)|$ is a scalar field over R.

15.4. If the temperature at points of space varies inversely as the square of the distance from a fixed plane, set up a function $f(P)$ which defines this scalar field.

15.5. Assuming Newton's law of gravitation, give a point function $f(P)$ which gives the magnitude of the force of attraction exerted on a unit particle by a point particle of mass M.

15.6. Set up the vector field function $V(P)$ for the force of attraction described in the preceding example.

15.7. Determine the scalar field function $f(P)$ whose level surfaces are confocal ellipsoids of revolution about the x-axis, the value of the field on each of the surfaces being proportional to the greater axis of the ellipsoid.

15.2 Directional derivative of $f(P)$.

Given a scalar function $F(x, y, z)$ the partial derivative of F with respect to x, $\partial F/\partial x$, is simply the derivative of F with respect to x which is obtained on the assumption that y and z are constants. Thus $\partial F/\partial x$ gives the rate of change of the function F with respect to x in the direction of increasing x. Thus we know how to obtain the rate of change of the function $f(P)$ with respect to a displacement along a direction parallel to one of the coördinate axes, that is, along a direction specified by one of the vectors $\boldsymbol{i}$, $\boldsymbol{j}$, or $\boldsymbol{k}$. We now consider the rate of change of the function with respect to a linear displacement in *any direction*.

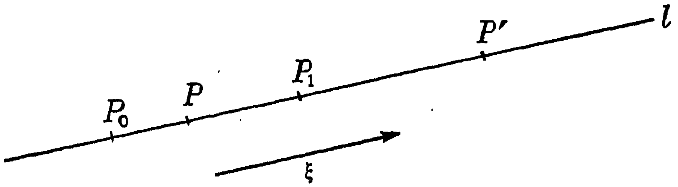

Fig. 28

Let $f(P)$ be a scalar field over R and let ξ be a unit vector. Through a fixed point P of R let the line l be constructed parallel to ξ (Fig. 28). The point P and the vector ξ form a coördinate system on l. Let P and P' be points on l whose coördinates are s and $s + \Delta s$, respectively, s being

arc length on l. If the quotient

$$\frac{f(P') - f(P)}{\Delta s}$$

has a limit as Δs approaches zero, the function $f(P)$ is said to admit at P a *derivative in the direction* $\boldsymbol{\xi}$. We write

$$\frac{df(P)}{ds} = \lim_{\Delta s \to 0} \frac{f(P') - f(P)}{\Delta s} \text{ with } \overrightarrow{PP'} = \Delta s \boldsymbol{\xi}.$$

We assume that at each point P of R the field $f(P)$ admits a derivative for every direction $\boldsymbol{\xi}$.

Since along the line l the scalar function $f(P)$ is a function of the real scalar variable s only, we have by the Law of the Mean of the differential calculus,

$$f(P') = f(P) + \Delta s \left(\frac{df(P)}{ds} \right)_{P_1},$$

where P_1 is a point on l between P and P'.

Exercises

15.8. Obtain the directional derivatives of $f(P) = 1/z^2$ at the point (1, 2, 3) for the directions

(1) $\boldsymbol{\xi} = \boldsymbol{k}$

(2) $\boldsymbol{\xi} = \cos \alpha \boldsymbol{i} + \sin \alpha \boldsymbol{j}$

(3) $\boldsymbol{\xi} = \cos \alpha \boldsymbol{i} + \cos \beta \boldsymbol{j} + \cos \gamma \boldsymbol{k}$,

where $\cos^2 \alpha + \cos^2 \beta + \cos^2 \gamma = 1$. In the last case show that there is a "cone of directions" along which $df(P)/ds$ has the same value.

15.9. If $f(P)$ stands for the function $F(x, y, z)$, show that $df(P)/ds$ in the direction $\boldsymbol{\xi} = \cos \alpha \boldsymbol{i} + \cos \beta \boldsymbol{j} + \cos \gamma \boldsymbol{k}$ is given by

$$\cos \alpha \frac{\partial F}{\partial x} + \cos \beta \frac{\partial F}{\partial y} + \cos \gamma \frac{\partial F}{\partial z}.$$

15.10. Show that, if the directional derivatives of $f(P)$ exist for the directions $\boldsymbol{i}$, $\boldsymbol{j}$, and $\boldsymbol{k}$, the directional derivative will exist for an arbitrary direction.

15.3 Gradient of a scalar field.

Consider the directional derivative of a scalar field $F(x, y, z)$ at a point P in the direction of the unit vector

ξ. Let C be a curve whose equation is

$$r = x(s)i + y(s)j + z(s)k$$

referred to its arc length s, which passes through P and whose unit tangent vector at that point is ξ. Then along C the field function F is a function of s only, and we have by the differential calculus

$$\frac{dF}{ds} = \frac{\partial F}{\partial x}\frac{dx}{ds} + \frac{\partial F}{\partial y}\frac{dy}{ds} + \frac{\partial F}{\partial z}\frac{dz}{ds}.$$

This may be written in the form

$$\frac{dF}{ds} = \left(\frac{dx}{ds}i + \frac{dy}{ds}j + \frac{dz}{ds}k\right)\cdot\left(\frac{\partial F}{\partial x}i + \frac{\partial F}{\partial y}j + \frac{\partial F}{\partial z}k\right).$$

The first vector on the right is the unit tangent vector of the curve C and, hence, at P is the vector ξ. The second vector on the right depends *only on the point P and not upon s or ξ*; it is called the *gradient of the scalar field F*. We write

$$\mathbf{grad}\ F = \frac{\partial F}{\partial x}i + \frac{\partial F}{\partial y}j + \frac{\partial F}{\partial z}k.$$

Hence, in terms of this concept the directional derivative is given by

$$\frac{dF}{ds} = \xi \cdot \mathbf{grad}\ F.$$

If ξ is tangential to a level surface F = constant, then

$$\frac{dF}{ds} = 0 = \xi \cdot \mathbf{grad}\ F.$$

Therefore, at each point P, **grad** F is *normal to the level surface* of the field which passes through P. Also, if ξ is in the direction of **grad** F, dF/ds is a non-negative number. Hence the vector **grad** F is (1) normal to the level surface and (2) *points in the direction of F increasing.* These two properties of **grad** F can be taken as defining the gradient. (Cf. *Juvet* (11), pp. 80–81.)

A curve which is orthogonal to each of a one-parameter family of level surfaces is called a *field curve* or *field line*, or sometimes a *line of flow*. Let dr represent a differential

displacement along a field line determined by the scalar field $F(x, y, z)$. Then $\boldsymbol{dr}$ and $\boldsymbol{grad}\ F$ are linearly dependent, and the differential equations of the field lines are

$$\frac{dx}{\dfrac{\partial f}{\partial x}} = \frac{dy}{\dfrac{\partial f}{\partial y}} = \frac{dz}{\dfrac{\partial f}{\partial z}}.$$

Exercises

15.11. Given the scalar field defined by

$$F(x, y, z) = \sqrt{x^2 + y^2 + z^2},$$

obtain $\boldsymbol{grad}\ F$ and verify that this vector at a point is normal to the level surface passing through that point.

15.12. If a scalar field is defined by $F(x, y, z) = g(r)$ where $r = \sqrt{x^2 + y^2 + z^2}$, show that

$$\boldsymbol{grad}\ F \times \boldsymbol{r} = \boldsymbol{0},$$

where $\boldsymbol{r}$ is the position vector of the point P at which $\boldsymbol{grad}\ F$ is evaluated.

15.13. The rate of change of $f(P)$ with respect to a linear displacement is a maximum for a displacement in the direction of $\boldsymbol{grad}\ f$.

15.14. Given P, $f(P)$, and λ where $0 < \lambda < |\boldsymbol{grad}\ f|$, show that there is a cone of directions $\boldsymbol{\xi}$ such that $df/ds = \lambda$.

15.15. A necessary condition that $f(P)$ have a maximum (or a minimum) at P_0 is that

$$(\boldsymbol{grad}\ f)_{P_0} = \boldsymbol{0}.$$

15.16. The field lines are determined by the differential equation

$$\boldsymbol{dr} \times \boldsymbol{grad}\ f = \boldsymbol{0}.$$

15.17. Let O; $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$ determine an affine coördinate system in which the coördinates of a point are denoted by (x_1, x_2, x_3). Let $F(x_1, x_2, x_3)$ be a given scalar field.

(1) Show that $\partial F/\partial x_1$, $\partial F/\partial x_2$, $\partial F/\partial x_3$ are the coefficients of a covariant vector field. (We write $\boldsymbol{grad}\ F$ for this vector.)

(2) dx_1, dx_2, dx_3 are the coefficients of a contravariant vector (which we write as $\boldsymbol{dr}$).

(3) Show that $dF = d\mathbf{r} \cdot \textit{\textbf{grad}}\, F$, and that dF is a scalar invariant with respect to the group of affine transformations.

(4) Show that $\textit{\textbf{grad}}\, F$ is orthogonal to the level surfaces of the scalar field.

15.18. If λ is a constant and f and g are scalar fields, show that:

(1) $\textit{\textbf{grad}}\,(f + g) = \textit{\textbf{grad}}\, f + \textit{\textbf{grad}}\, g$

(2) $\textit{\textbf{grad}}\,(\lambda f) = \lambda\, \textit{\textbf{grad}}\, f$

(3) $\textit{\textbf{grad}}\,(fg) = f\, \textit{\textbf{grad}}\, g + g\, \textit{\textbf{grad}}\, f.$

CHAPTER III

Integral Calculus of Vectors

§16. Definite Integrals

References: *Goursat* (34), pp. 140–151; *Gibson* (33), Chapter IX; *Osgood* (45), Chapters II and XI; *Courant, R.*, Differential and Integral Calculus, Vol. II, tr. by E. J. McShane, New York, Nordemann, 1938.

In beginning our study of the integral calculus of vectors we recall the following definition of a definite integral.

Let $f(x)$ be a real function of a real variable x defined on the interval $a \leqq x \leqq b$. Let δ be an arbitrarily assigned positive number and let the interval $[a, b]$ be divided into a number of subintervals, the division points being

$$x_0, x_1, x_2, \cdots, x_{n-1}, x_n,$$

where $x_0 = a$, $x_n = b$, and which satisfy the inequalities $x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n$, and such that the length of each subinterval is at most equal to δ. Denote the length of the ith subinterval by Δx_i; that is, set $\Delta x_i = x_i - x_{i-1}$. Let $\bar{x}_i$ be any value of x interior to or on the boundary of the ith subinterval, and form the sum

$$f(\bar{x}_1)\Delta x_1 + f(\bar{x}_2)\Delta x_2 + \cdots + f(\bar{x}_n)\Delta x_n = \sum_{i=1}^{n} f(x_i)\Delta x_i.$$

If the limit of this sum exists as δ approaches zero, and has a value which is independent of the mode of division into subintervals, and of the choice of $\bar{x}_i$ in each subinterval, it is called *the definite integral in the sense of Riemann of the function $f(x)$ from a to b.* The definite integral is denoted by the symbol

$$\int_a^b f(x)dx.$$

The so-called Fundamental Theorem of the integral calculus may be stated as follows. If the definite integral exists and $\varphi(x)$ is defined by

$$\varphi(x) = \int_a^x f(x)dx,$$

then $d\varphi(x)/dx = f(x)$ for each value of x at which $f(x)$ is continuous.

16.1 Line integrals.

References: *Goursat* (34), pp. 184–189; *Gibson* (33), pp. 296–300; *Osgood* (45), Chapter XI; *Picard* (47), Vol. 1, Chapter III.

Let C be an oriented curve whose equation is

$$\boldsymbol{r} = \boldsymbol{r}(t).$$

We consider the line integrals

$$\int_C d\boldsymbol{r}f(P); \int_C d\boldsymbol{r}\cdot \boldsymbol{V}(P); \int_C d\boldsymbol{r}\times \boldsymbol{V}(P),$$

defined as limits of the appropriate sums, where $f(P)$ and $\boldsymbol{V}(P)$ are given fields defined over the points of C under consideration. The first and third integrals yield a vector; the second yields a scalar. Diagrammatically,

$$\int_C d\boldsymbol{r}\begin{cases}\text{scalar}\\ \cdot\ \text{vector}\\ \times\ \text{vector}\end{cases} = \begin{cases}\text{vector}\\ \text{scalar}\\ \text{vector.}\end{cases}$$

Along a given curve

$$C:\quad \boldsymbol{r}(t) = x(t)\boldsymbol{i} + y(t)\boldsymbol{j} + z(t)\boldsymbol{k}$$

all the functions involved in the integrand are of course known in terms of the variable (or parameter) t. When the integrand has been completely expressed in terms of this variable, the line integral becomes an ordinary definite integral and is evaluated in the usual manner. For given fields and given end points A and B, the role which the curve plays in these integrals is in determining just what function of t the integrand becomes. The evaluation of

any one of the integrals may be carried out in terms of any convenient variable or variables. Thus, considering the first one

$$\int_A^B d\boldsymbol{r} f(P) = \int_{t_1}^{t_2} \frac{d\boldsymbol{r}(t)}{dt} F[x(t), y(t), z(t)]dt,$$

where t_1 and t_2 correspond to the points A and B, respectively. However, we may express $d\boldsymbol{r}$ as

$$d\boldsymbol{r} = \boldsymbol{i}\,dx + \boldsymbol{j}\,dy + \boldsymbol{k}\,dz,$$

and the integral in the form

$$\int_A^B (\boldsymbol{i}\,dx + \boldsymbol{j}\,dy + \boldsymbol{k}\,dz)F(x, y, z)$$
$$= \int_{x_1}^{x_2} \boldsymbol{i}F(x, y, z)dx + \int_{y_1}^{y_2} \boldsymbol{j}F(x, y, z)dy + \int_{z_1}^{z_2} \boldsymbol{k}F(x, y, z)dz,$$

where the limits of integration are the corresponding coördinates of the end points A and B. In the first of these integrals y and z must be replaced by the expressions which they assume in terms of x for points on the curve C; this is where the curve enters into the problem. Similar remarks apply to the remaining two integrals.

Since we are supposing a system of constant base vectors $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$, it follows from the definition of a definite integral, and the property that scalar multiplication of vectors is distributive with respect to addition of scalars, that

$$\int_{x_1}^{x_2} \boldsymbol{i}F(x, y, z)dx = \boldsymbol{i}\int_{x_1}^{x_2} F(x, y, z)dx.$$

As an illustrative example, consider the integral

$$\int_C d\boldsymbol{r} \cdot \boldsymbol{V}$$

along the curve

$$C: \quad \boldsymbol{r} = t\boldsymbol{i} + \frac{t^2}{2}\boldsymbol{j} + 2t^3\boldsymbol{k}$$

from $A(0, 0, 0)$ to $B(1, \frac{1}{2}, 2)$, where the vector field is given by

$$\boldsymbol{V}(P) = y\boldsymbol{i} - x^2\boldsymbol{j} + xyz\boldsymbol{k}.$$

Since r is the position vector of points on the curve C, the coördinates (x, y, z) of any such point are given in terms of t by the equations

$$x = t,\ y = \frac{t^2}{2},\ z = 2t^3.$$

The point A corresponds to $t = 0$, and the point B to $t = 1$. Hence

$$\begin{aligned}\int_A^B dr \cdot V &= \int_0^1 dt(i + tj + 6t^2k) \cdot \left(\frac{t^2}{2}i - t^2j + t^6k\right)\\ &= \int_0^1 \left(\frac{t^2}{2} - t^3 + 6t^8\right)dt = \frac{7}{12}.\end{aligned}$$

Computing this integral by the other method, we have

$$\begin{aligned}\int_A^B dr \cdot V &= \int_A^B (idx + jdy + kdz) \cdot (yi - x^2j + xyzk)\\ &= \int_A^B (ydx - x^2dy + xyzdz)\\ &= \int_0^1 \frac{x^2}{2}dx - \int_0^{1/2} 2ydy + \int_0^2 \frac{z^2}{4}dz\\ &= \frac{7}{12}.\end{aligned}$$

In this example, if V is interpreted as a force field, the result measures the work done in transporting a unit particle along the curve from the point A to B. Generally, the value of a line integral depends not only upon the end points but also on the path. Thus the curve

$$r = \frac{t}{2}i + \frac{t}{4}j + \frac{t^2}{2}k$$

contains the point A for $t = 0$ and the point B for $t = 2$. However, in this case

$$\begin{aligned}\int_A^B dr \cdot V &= \int_0^2 dt\left(\frac{i}{2} + \frac{j}{4} + tk\right) \cdot \left(\frac{t}{4}i - \frac{t^2}{4}j + \frac{t^4}{16}k\right)\\ &= \int_0^2 \left(\frac{t}{8} - \frac{t^2}{16} + \frac{t^5}{16}\right)dt = \frac{3}{4}.\end{aligned}$$

Evaluating by the other method, we have

$$\begin{aligned}\int_A^B dr \cdot V &= \int_A^B (idx + jdy + kdz) \cdot (yi - x^2j + xyzk) \\ &= \int_A^B (ydx - x^2dy + xyzdz) \\ &= \int_0^1 \frac{x}{2}dx - \int_0^{\frac{1}{2}} 4y^2dy + \int_0^2 \frac{z^2}{4}dz \\ &= \frac{3}{4}.\end{aligned}$$

Exercises

16.1. Identify the "ordinary integral"

$$\int_{x_1}^{x_2} f(x)dx$$

with one of the above line integrals. Give a geometric interpretation for the special case $f(x) \equiv 1$.

16.2. Give physical or geometric problems which afford an illustration for each of the three types of line integrals.

16.3. Evaluate each of the three types of line integrals along each of the curves

$$\begin{aligned}C_1&: \ r(t) = ti + tj + tk \\ C_2&: \ r(t) = ti + t^2j + t^3k\end{aligned}$$

from the point (0, 0, 0) to the point (1, 1, 1), given that

$$\begin{aligned}f(P) &= x^2 - yz^2 \\ V(P) &= xyi - z^2j + xyzk.\end{aligned}$$

Also, evaluate

$$\int dr \cdot \mathrm{grad}\, f$$

along the same two curves between the same points.

16.4. The arc length s of the curve $C: r = r(t)$ from the point given by $t = t_1$ to the point given by t is

$$s = \int_{t_1}^{t} dr \cdot \xi_1,$$

where ξ_1 is the unit tangent vector of the curve.

16.2 Surface integrals.

References: *Goursat* (34), p. 256 *ff.*; *Gibson* (33), pp. 312–323; *Osgood* (45), Chapter III.

Let Σ be a *two-sided surface.* We suppose that the surface has a unique normal at each point under consideration. At an arbitrary point P on the surface let a positive direction be assigned to the normal. The supposition that the surface is two-sided means that if the point P, accompanied by its oriented normal, traverses any closed curve on the surface, upon its return to its original position the normal will be oriented as initially. This is not true, for instance, on the Möbius strip. (Cf. *Veblen-Young* (51), II, p. 67.)

We shall assume that any surface under consideration is orientable and that it has been oriented by a definite choice of its positive normal. Likewise, it will be understood that any curve under consideration has been oriented.

We shall be concerned with only two types of orientable surfaces:

(1) A simply connected surface with a boundary consisting of a closed curve. In this case we take the orientations of the curve C and the surface S such that the tangent vector of C, its inward-pointing normal (tangent to the surface), and the normal to the surface, in the order named, constitute a right-handed triple.

(2) The surface which bounds a finite, simply connected region of space. In this case the outward-pointing normal will be chosen as the normal to the surface.

Let $d\sigma$ be an element of area of Σ, and let $\boldsymbol{\xi}$ be the unit normal vector to the surface. It is convenient to introduce the surface element vector $\boldsymbol{d\hat{\sigma}}$, defined by

$$\boldsymbol{d\hat{\sigma}} = d\sigma\boldsymbol{\xi}.$$

We consider the surface integrals

$$\int_\Sigma \boldsymbol{d\hat{\sigma}} f(P),\ \int_\Sigma \boldsymbol{d\hat{\sigma}} \cdot V(P),\ \int_\Sigma \boldsymbol{d\hat{\sigma}} \times V(P).$$

defined as limits of the appropriate sums. Diagrammatically,

$$\int_{\Sigma} d\boldsymbol{\sigma} \begin{cases} \text{scalar} \\ \cdot \text{ vector} \\ \times \text{ vector} \end{cases} = \begin{cases} \text{vector} \\ \text{scalar} \\ \text{vector.} \end{cases}$$

Let the vector equation of the surface Σ be

$$\boldsymbol{r} = \boldsymbol{r}(u, v).$$

Then, at any point P of the surface, $f(P)$ and $\boldsymbol{V}(P)$ are known in terms of the scalar variables u and v. Also, as we have seen $d\boldsymbol{\sigma}$ can be expressed in terms of these variables, viz.:

$$d\boldsymbol{\sigma} = (\boldsymbol{r}_u \times \boldsymbol{r}_v)du\,dv.$$

It is proved in works on analysis that if the double integral

$$\int F(u, v)d\sigma$$

exists, then the repeated integral

$$\int dv \int F(u, v) g(u, v) du = \int du \int F(u, v) g(u, v) dv,$$

where $d\sigma = g(u, v)du\,dv$, also exists and has the same value. See *Goursat* (34) p. 254, or *E. W. Hobson*, Theory of Functions of a Real Variable, Cambridge (1927), Vol. I, p. 510.

As an illustration of one of these integrals we consider

$$\int_{\Sigma} d\boldsymbol{\sigma} \times \boldsymbol{V}(P)$$

where Σ is the closed surface consisting of Σ_1, the hemispherical surface above the x, y-plane of a sphere with center at the origin and radius equal to one, and Σ_2, the surface of the x, y-plane interior to and including the unit circle with center at the origin. For the vector field we take

$$\boldsymbol{V}(P) = y\boldsymbol{i} - x^2\boldsymbol{j} + xz\,\boldsymbol{k}.$$

The position vector $\boldsymbol{r}$ of points on the sphere is

$$\boldsymbol{r} = \cos u \cos v\boldsymbol{i} + \cos u \sin v\boldsymbol{j} + \sin u\,\boldsymbol{k},$$

where u is latitude and v is longitude. The surface element for the spherical surface is given by $d\sigma = \cos u\, du\, dv$, and hence the outward-pointing surface element vector $\boldsymbol{d\sigma}$ is given by

$$\begin{aligned} \boldsymbol{d\sigma} &= r \cos u\, du\, dv \\ &= (\cos^2 u \cos v\boldsymbol{i} + \cos^2 u \sin v\boldsymbol{j} + \sin u \cos u\,\boldsymbol{k})du\,dv. \end{aligned}$$

At points on the spherical surface the vector field $\boldsymbol{V}(P)$ is given in terms of u and v by replacing x, y, and z by their expressions in terms of u and v. Hence on the spherical surface

$$\boldsymbol{V}(P) = \cos u \sin v\boldsymbol{i} - \cos^2 u \cos^2 v\boldsymbol{j} + \sin u \cos u \cos v\boldsymbol{k}.$$

A direct computation now gives $\boldsymbol{d\sigma} \times \boldsymbol{V}(P)$ for points of Σ_1, and we have

$$\begin{aligned} \int_{\Sigma_1} \boldsymbol{d\sigma} \times \boldsymbol{V}(P) &= \boldsymbol{i}\int_0^{2\pi}\int_0^{\frac{\pi}{2}}(\cos^3 u \sin u \cos v \sin v + \\ &\qquad\qquad \cos^3 u \sin u \cos^2 v)du\,dv \\ &\quad + \boldsymbol{j}\int_0^{2\pi}\int_0^{\frac{\pi}{2}}(\cos^2 u \sin u \sin v - \cos^3 u \sin u \cos^2 v)du\,dv \\ &\quad + \boldsymbol{k}\int_0^{2\pi}\int_0^{\frac{\pi}{2}}(-\cos^4 u \cos^3 v - \cos^3 u \sin^2 v)du\,dv \\ &= \frac{\pi}{4}\boldsymbol{i} - \frac{\pi}{4}\boldsymbol{j} - \frac{2\pi}{3}\boldsymbol{k}. \end{aligned}$$

On the surface Σ_2 it will be convenient to use polar coördinates ρ, θ. Then the element of area is given by $\rho d\rho d\theta$, and hence

$$\boldsymbol{d\sigma} = -\rho d\rho d\theta \boldsymbol{k}.$$

In terms of these variables

$$\begin{aligned} x &= \rho \cos \theta \\ y &= \rho \sin \theta, \end{aligned}$$

and hence for points of Σ_2

$$\boldsymbol{V}(P) = (\rho \sin \theta)\boldsymbol{i} - (\rho^2 \cos^2 \theta)\boldsymbol{j}.$$

Then

$$\int_{\Sigma_2} d\mathbf{\sigma} \times V(P) = i\int_0^{2\pi}\int_0^1 - \rho^3 \cos^2 \theta d\rho d\theta + j\int_0^{2\pi}\int_0^1 - \rho^2 \sin \theta d\rho d\theta$$

$$= -\frac{\pi}{4}i.$$

Combining these results, we have

$$\int_{\Sigma} d\mathbf{\sigma} \times V(P) = -\frac{\pi}{4}j - \frac{2\pi}{3}k.$$

Note.—The discerning reader will observe that the surface Σ violates somewhat our hypothesis, inasmuch as it has no well-defined normal along the curve of intersection of the sphere and the x, y-plane. However, the problem is typical of those which may arise, and the result in this instance can be justified by the following device. In the case of an integral over Σ_1, let it be defined by

$$\lim_{\alpha \rightharpoonup 0} \int_{\alpha}^{\frac{\pi}{2}}\int_0^{2\pi} \varphi(u, v) dv\, du,$$

where the symbol $\alpha \rightharpoonup 0$ means that α is to approach 0 through positive values, and for an integral over Σ_2 we take as its value

$$\lim_{r \to 1} \int_0^r\int_0^{2\pi} F(\rho, \theta) d\theta\, d\rho,$$

where r is to approach 1 from below. For an adequate treatment of integrals over surfaces, reference may be made to *Kellogg* (44), Chapter IV.

16.3 Volume integrals.

Let T be a closed region of space, and let $d\tau$ be an element of volume. We consider the *volume integrals*

$$\int_T d\tau\, f(P), \quad \int_T d\tau\, V(P)$$

defined as limits of the appropriate sums. Diagrammatically,

$$\int_T d\tau \begin{cases} \text{scalar} \\ \text{vector} \end{cases} = \begin{cases} \text{scalar} \\ \text{vector.} \end{cases}$$

As an example, we calculate the value of

$$\int_T d\tau\, V(P),$$

where

$$V(P) = y\boldsymbol{i} - x^2\boldsymbol{j} + xz\boldsymbol{k}$$

and T is the region bounded by the right circular cylinder of radius 1 whose axis is the z-axis and whose bases are the planes $z = 0$ and $z = 2$. Using cylindrical coördinates ρ, θ, z,

$$\begin{aligned} d\tau &= \rho d\rho d\theta dz \\ \int_T d\tau V(P) &= \int_0^{2\pi}\int_0^2\int_0^1 \tfrac{1}{2}\rho\,(\rho \sin\theta\boldsymbol{i} - \rho^2\cos^2\theta\boldsymbol{j} + \rho\cos\theta z\boldsymbol{k})d\rho dz d\theta \\ &= -\tfrac{1}{2}\boldsymbol{j}\int_0^2\int_0^1\int_0^{2\pi} \rho^3\cos^2\theta d\theta\, d\rho dz \\ &= -\frac{\pi}{4}\boldsymbol{j}. \end{aligned}$$

Exercises

16.5. Give a formal definition of one of the above surface integrals.

16.6. Give physical or geometric problems which afford illustrations of the surface and volume integrals under discussion.

16.7. Given the equation of the right circular cylinder of radius R in the form

$$\boldsymbol{r} = R\cos u\boldsymbol{i} + R\sin u\boldsymbol{j} + v\boldsymbol{k},$$

(1) Compute the surface element vector $d\boldsymbol{\sigma}$, where the positive normal is taken as pointing outward.

(2) Determine by integration the area of the curved surface of height h.

(3) Given a pressure function $f(P)$ proportional to the depth below the upper face, compute the force vector on the convex surface of the half cylinder of height h lying to the right of the $\boldsymbol{i}$, $\boldsymbol{k}$-plane. Note that the force vector is the same as though the surface of the half cylinder were replaced by the $\boldsymbol{i}$, $\boldsymbol{k}$-diameter plane.

(4) Compute the force vector on the half cylinder under the conditions specified above, but including the force on the lower base.

16.8. The integral $\int_C d\mathbf{r} \cdot \mathbf{V}(P)$ is called the *circulation of* $\mathbf{V}$ *along* C; the integral $\int_\Sigma d\boldsymbol{\sigma} \cdot \mathbf{V}(P)$ is called the *flux of* $\mathbf{V}$ *across* Σ. Show that the arc length of C is the circulation of the unit tangent vector along C, and that the area of Σ is the flux of the unit normal vector across Σ.

16.9. If $\mathbf{V}$ is a constant vector field, the flux of $\mathbf{V}$ across a closed surface is zero. Give a physical illustration of the meaning of this theorem.

16.10. Suppose an incompressible fluid of density ρ is flowing in a circular pipe of radius R such that the magnitude of the velocity vector is $1/(1 + r^2)$, where r is the distance from the center of the pipe to the particle. Obtain the momentum vector of the liquid in a section of the pipe h units long. We suppose the section to be straight.

16.4 Notion of solid angle.

A geometric concept of importance is that of the *solid angle* made by a surface Σ when viewed from a point O, not on Σ. Let P be a point in an element of surface $d\sigma$ of Σ, and let

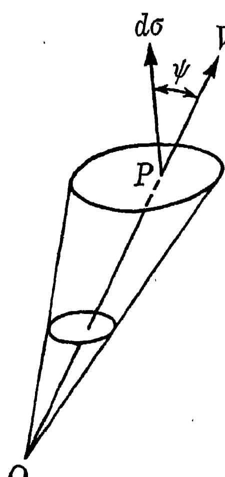

Fig. 29

$$\mathbf{r} = \overrightarrow{OP}.$$

We define a vector field

$$\mathbf{V}(P) = \frac{\mathbf{r}}{r^3},$$

where r is the magnitude of $\mathbf{r}$. Let ψ be the angle between $\mathbf{V}$ and $d\boldsymbol{\sigma}$ at P. Then

$$\mathbf{V} \cdot d\boldsymbol{\sigma} = |\mathbf{V}| d\sigma',$$

where $d\sigma'$ is the orthogonal projection of the area $d\sigma$ on a plane through P which is perpendicular to $\mathbf{V}$. Except for infinitesimals of higher order, $d\sigma'$ may equally well be interpreted as the area of the central projection from O of $d\sigma$ on the sphere with center at O and passing through P. Now the length of $\mathbf{V}$

is $1/r^2$, and $|Vd\sigma'|$ is the area, except for infinitesimals of higher order, of the sphere of radius 1 intercepted by the cone, with vertex at O, which projects $d\sigma$ on the sphere. For, if $d\sigma''$ denotes the area on the unit sphere

$$\frac{d\sigma''}{1} = \frac{d\sigma'}{r^2}$$

and

$$|Vd\sigma'| = \frac{|d\sigma'|}{r^2}.$$

The number $|Vd\sigma'|$ is called the *solid angle* of $d\sigma$ from O. The angle is said to be positive if $\overrightarrow{OP}$ enters Σ from the negative side as indicated in the figure. In the formula $|V| \cos \psi d\sigma$, the factor $\cos \psi$ takes care of the proper sign. The solid angle of the surface viewed from O is defined by the integral

$$\int_\Sigma d\mathbf{\hat{\sigma}} \cdot \frac{\mathbf{r}}{r^3},$$

$\mathbf{r}$ being the position vector $\overrightarrow{OP}$ of points P of Σ.

Exercise

16.11. Establish the following, known as Gauss's Theorem: If Σ is a simply connected closed surface, the solid angle subtended by Σ from a point O has the value 4π, 2π, 0 according as O is interior, on, or exterior to the surface Σ. Note that this is a scalar point function which is discontinuous at every point of the surface Σ.

§17. Differential Operators

17.1 Gradient, divergence, curl.

References: *Juvet* (11), Chapter IV; *Wills* (23), Chapter III.

Let $f(P)$ and $V(P)$ define fields over a region R. Let P be any point interior to R, and let Σ be any closed surface (with suitable continuity properties of course) situated in

R and containing P as an interior point. Consider the integral

$$\frac{\int_{\Sigma} d\delta f(Q)}{\tau},$$

where $d\delta$ is the outward-pointing normal and τ is the volume bounded by Σ. If the limit of this quotient exists as "points of Σ approach P," it defines a vector at P. Assuming the limits to exist, we adopt the following *definitions:*

$$\textbf{gradient } f = \lim_{\text{points of } \Sigma \to P} \frac{\int_{\Sigma} d\delta f(Q)}{\tau}$$

$$\textbf{divergence } V = \lim_{\text{points of } \Sigma \to P} \frac{\int_{\Sigma} d\delta \cdot V(Q)}{\tau}$$

$$\textbf{curl } V = \lim_{\text{points of } \Sigma \to P} \frac{\int_{\Sigma} d\delta \times V(Q)}{\tau}$$

Since the concept of limit demands that these limits, if they exist, shall have values independent of the shrinking surface Σ, we may obtain the values of the limits by con-

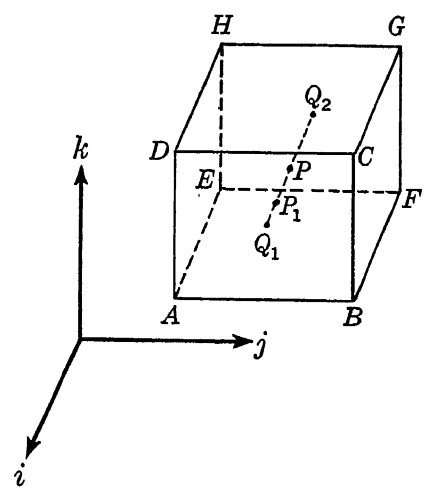

Fig. 30

sidering a particular surface. Making use of this principle, we now derive the expression for ***grad*** f in terms of an $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ system by taking Σ as a rectangular parallelepiped with its center at P and with its faces parallel to the coördinate planes.

Let P be the point (x, y, z) and let the dimensions of the parallelepiped be dx, dy, and dz. Suppose $f(P) = F(x, y, z)$. By the Law of the Mean, if Q_1 is the center of the face $ABCD$,

$$f(Q_1) = F\left(x + \frac{dx}{2}, y, z\right) = F(x, y, z) + \left(\frac{\partial F}{\partial x}\right)_{P_1}\frac{dx}{2},$$

where P_1 is a point between P and Q_1. On the face $ABCD$,

$$\begin{aligned} d\boldsymbol{\sigma} f(Q_1) &= \boldsymbol{i}\, dy\, dz f(Q_1) \\ &= \boldsymbol{i}\left\{F(P) + \left(\frac{\partial F}{\partial x}\right)_{P_1}\frac{dx}{2}\right\}dy\, dz. \end{aligned}$$

But on the opposite face $EFGH$,

$$\begin{aligned} d\boldsymbol{\sigma} f(Q_2) &= -\boldsymbol{i}\, dy\, dz f(Q_2) \\ &= -\boldsymbol{i}\left\{F(P) + \left(\frac{\partial F}{\partial x}\right)_{P_2}\left(-\frac{dx}{2}\right)\right\}dy\, dz, \end{aligned}$$

where Q_2 is the center of the face and P_2 is a point between P and Q_2. Hence the surface integral over these two faces of the parallelepiped gives

$$\frac{\boldsymbol{i}}{2}\left\{\left(\frac{\partial F}{\partial x}\right)_{P_1} + \left(\frac{\partial F}{\partial x}\right)_{P_2}\right\}dx\, dy\, dz.$$

Dividing this by the volume $d\tau = dx\, dy\, dz$ and passing to the limit we obtain

$$\boldsymbol{i}\left(\frac{\partial F}{\partial x}\right)_P.$$

Proceeding similarly for the other two pairs of opposite faces, we have

$$\boldsymbol{grad}\, f = \frac{\partial F}{\partial x}\boldsymbol{i} + \frac{\partial F}{\partial y}\boldsymbol{j} + \frac{\partial F}{\partial z}\boldsymbol{k}.$$

This result identifies the gradient as here defined with the quantity previously introduced under the same name.

A more satisfying approach to the expression for the gradient is furnished by selecting the surface which is to close down on P, one which has a well-defined normal at each point. Let, then, Σ be the sphere of radius r with the point P as its center. If P has coördinates (x, y, z), the position vector of a point Q on the sphere is

$$\overrightarrow{OQ} = (x + r\cos u \cos v)\boldsymbol{i} + (y + r\cos u \sin v)\boldsymbol{j} + (z + r\sin u)\boldsymbol{k}.$$

The surface element vector is

$$d\boldsymbol{\sigma} = r^2 (\cos u \cos v\boldsymbol{i} + \cos u \sin v\boldsymbol{j} + \sin u\boldsymbol{k}) \cos u\, du\, dv.$$

We expand $f(Q)$ by Taylor's theorem for a given u, v in terms of r:

$$f(Q) = f(P) + \left(\frac{\partial f}{\partial x}\cos u \cos v + \frac{\partial f}{\partial y}\cos u \sin v + \frac{\partial f}{\partial z}\sin u\right)r + R(r),$$

where the partial derivatives are of course evaluated at P and where the remainder term $R(r)$ is an infinitesimal of higher order with respect to r. Hence

$$\int_\Sigma d\boldsymbol{\sigma} f(Q) = \int_\Sigma d\boldsymbol{\sigma} f(P) + r^3\left\{\int_0^{2\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\left[\frac{\partial f}{\partial x}\cos^3 u \cos^2 v\boldsymbol{i} + \frac{\partial f}{\partial y}\cos^3 u \sin^2 v\boldsymbol{j} + \frac{\partial f}{\partial z}\sin^2 u \cos u\boldsymbol{k}\right]du\, dv\right\} + \int_\Sigma d\boldsymbol{\sigma} R(r).$$

The first integral on the right vanishes, and upon evaluating the repeated integrals, one obtains

$$\int_\Sigma d\boldsymbol{\sigma} f(Q) = \frac{4}{3}\pi r^3\left\{\frac{\partial f}{\partial x}\boldsymbol{i} + \frac{\partial f}{\partial y}\boldsymbol{j} + \frac{\partial f}{\partial z}\boldsymbol{k} + \boldsymbol{\varepsilon}(r)\right\},$$

where the vector $\boldsymbol{\varepsilon}(r)$ approaches zero as r approaches zero. Now, dividing this expression by the volume of the sphere and passing to the limit, we obtain the gradient of f in the same form as before.

Exercises

17.1. Given the vector field

$$V(P) = X(x, y, z)\boldsymbol{i} + Y(x, y, z)\boldsymbol{j} + Z(x, y, z)\boldsymbol{k},$$

show that

$$\text{divergence } V = \frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y} + \frac{\partial Z}{\partial z}.$$

and that

$$\text{curl } V = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ X & Y & Z \end{vmatrix}.$$

17.2. Given the vector field

$$V(P) = x \cos y\,\boldsymbol{i} + y^2 z\,\boldsymbol{j} - xyz\,\boldsymbol{k},$$

compute div V and curl V.

17.3. A necessary and sufficient condition that grad f vanish identically is that f = constant.

17.4. If $V(P) = \text{grad } f(P)$, show that curl $V \equiv 0$.

17.5. Show that div (curl V) $\equiv 0$.

17.2 Differential operators.

In this connection it is convenient to introduce the operator "del," denoted by ∇ and defined by

$$\nabla \equiv \boldsymbol{i}\frac{\partial}{\partial x} + \boldsymbol{j}\frac{\partial}{\partial y} + \boldsymbol{k}\frac{\partial}{\partial z}.$$

In terms of this symbol we may write

$$\begin{aligned} \text{grad } f &= \nabla f \\ \text{div } V &= \nabla \cdot V \\ \text{curl } V &= \nabla \times V. \end{aligned}$$

We also introduce two additional scalar operators, $V \cdot \nabla$ and $\nabla \cdot \nabla$. Let us consider the latter one first.

$$\begin{aligned} \nabla \cdot \nabla &= \left(\boldsymbol{i}\frac{\partial}{\partial x} + \boldsymbol{j}\frac{\partial}{\partial y} + \boldsymbol{k}\frac{\partial}{\partial z}\right) \cdot \left(\boldsymbol{i}\frac{\partial}{\partial x} + \boldsymbol{j}\frac{\partial}{\partial y} + \boldsymbol{k}\frac{\partial}{\partial z}\right) \\ &= \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}. \end{aligned}$$

In place of $\nabla \cdot \nabla$ we employ the more customary notation ∇^2. The operator ∇^2 is called the *Laplacian*, and the equation $\nabla^2 f = 0$ is known as *Laplace's Equation.*

We have seen that

$$df = d\mathbf{r} \cdot \text{grad}\, f,$$

which may be written

$$df = (\mathbf{s} \cdot \nabla) f\, ds,$$

where $d\mathbf{r} = \mathbf{s}\, ds$, $\mathbf{s}$ being a unit vector. Thus the operator $(\mathbf{s} \cdot \nabla)$, where $\mathbf{s}$ is a unit vector, when applied to a scalar field f, yields a number which is the directional derivative df/ds in the direction $\mathbf{s}$. That is, $\mathbf{s}$ being a unit vector, we have

$$\mathbf{s} \cdot (\nabla f) = (\mathbf{s} \cdot \nabla) f.$$

We see from the definition of the gradient of f that we can express the result of $(\mathbf{s} \cdot \nabla)$ operating on $f(P)$ in the form

$$(\mathbf{s} \cdot \nabla) f(P) = \lim_{\text{points of } \Sigma \to P} \frac{\int_\Sigma \mathbf{s} \cdot d\boldsymbol{\sigma} f(Q)}{\tau}$$

where now $\mathbf{s}$ stands for a constant field of unit vectors.

Let $\mathbf{V}(P)$ be a vector field. We define the "directional derivative of $\mathbf{V}(P)$ in the direction $\mathbf{s}$" by

$$\frac{d\mathbf{V}}{ds} = \lim_{\Delta s \to 0} \frac{\mathbf{V}(P') - \mathbf{V}(P)}{\Delta s},$$

where $\overrightarrow{PP'} = \mathbf{s}\Delta s$.

We now verify that

$$\frac{d\mathbf{V}}{ds} = (\mathbf{s} \cdot \nabla) \mathbf{V},$$

where $(\mathbf{s} \cdot \nabla)\mathbf{V}$ is defined by

$$(\mathbf{s} \cdot \nabla)\mathbf{V}(P) = \lim_{\text{points of } \Sigma \to P} \frac{\int_\Sigma (\mathbf{s} \cdot d\boldsymbol{\sigma}) \mathbf{V}(Q),}{\tau},$$

$\mathbf{s}$ being a constant unit vector field. For the evaluation of the integral, we select Σ as a right circular cylinder with

PP' as axis, and with P an interior point. Then at each point of the curved surface $\boldsymbol{s} \cdot d\boldsymbol{\sigma} = 0$. Let α be the area of the base of the cylinder. Then

$$\lim_{\text{points of } \Sigma \to P} \frac{\int_\Sigma (\boldsymbol{s} \cdot d\boldsymbol{\sigma}) V(Q)}{\tau} = \lim_{\text{points of } \Sigma \to P} \frac{V(P')\alpha - V(P)\alpha}{\alpha |\overrightarrow{PP'}|}$$

$$= \lim_{\Delta s \to 0} \frac{V(P') - V(P)}{\Delta s}.$$

Hence

$$\frac{dV}{ds} = (\boldsymbol{s} \cdot \nabla) V.$$

In terms of an $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ system

$$\boldsymbol{s} \cdot \nabla = s_1 \frac{\partial}{\partial x} + s_2 \frac{\partial}{\partial y} + s_3 \frac{\partial}{\partial z},$$

where s_1, s_2, and s_3 are the direction cosines of $\boldsymbol{s}$.

Similarly one can define the operators

$$(\boldsymbol{s} \times \nabla) \cdot V(P) = \lim_{\text{points of } \Sigma \to P} \frac{\int_\Sigma (s \times d\boldsymbol{\sigma}) \cdot V(Q)}{\tau},$$

and

$$(\boldsymbol{s} \times \nabla) \times V(P) = \lim_{\text{points of } \Sigma \to P} \frac{\int_\Sigma (s \times d\boldsymbol{\sigma}) \times V(Q)}{\tau}.$$

From the calculus, if λ is a parameter independent of x and

$$\varphi(\lambda) = \int_{x_1}^{x_2} f(\lambda, x) dx,$$

the limits of integration being independent of λ, then

$$\frac{d\varphi}{d\lambda} = \int_{x_1}^{x_2} \frac{\partial f(\lambda, x)}{\partial \lambda} dx.$$

We now establish the analogue of this theorem for the operator del.

Let $f(P, Q)$ be a scalar function of two independent points. Then $g(P)$ is a scalar point function, $g(P)$ being

defined by

$$g(P) = \int_\Sigma f(P, Q)d\sigma,$$

where during the integration P is fixed and Q varies over a surface Σ. Let P' be a point near to P such that $\overrightarrow{PP'} = \Delta s\,\mathbf{s}$. Then

$$\frac{g(P') - g(P)}{\Delta s} = \int_\Sigma \frac{f(P', Q) - f(P, Q)}{\Delta s} d\sigma.$$

Now if $f(P, Q)$ regarded as a function of P has a gradient

$$f(P', Q) - f(P, Q) = \Delta s \left(\frac{df(P, Q)}{ds}\right)_{P_1},$$

where P_1 is between P and P'. Then in the limit, if the gradient is continuous,

$$\frac{dg(P)}{ds} = \int_\Sigma \left(\frac{df}{ds}\right)_P d\sigma.$$

That is,

$$\mathbf{s} \cdot \operatorname{grad} g(P) = \int_\Sigma \mathbf{s} \cdot \operatorname{grad}_P f(P, Q)d\sigma,$$

and since this is true for $\mathbf{s}$ arbitrary, we conclude that

$$\nabla g(P) = \int_\Sigma \nabla_P f(P, Q)d\sigma,$$

where ∇_P is used to indicate that ∇ operates on $f(P, Q)$ regarded as a function of P only. The result which has just been established in the case of a surface integral clearly holds also in the case of an integration along a curve or over a volume, the limits of integration of course being independent of P.

As a working rule one makes use of the fact that the operator ∇ behaves somewhat as a vector, except of course that in the notation it must precede the quantity on which it operates. In any case in which the meaning may be in doubt, one can usually resort to the limit definitions, as has been done above in establishing most of the results.

As an example, consider

$$\nabla \cdot (\boldsymbol{u} \times \boldsymbol{v}).$$

Since ∇ is a differential operator, we have

$$\begin{aligned}\nabla \cdot (\boldsymbol{u} \times \boldsymbol{v}) &= \nabla \cdot (\underset{\uparrow}{\boldsymbol{u}} \times \boldsymbol{v}) + \nabla \cdot (\boldsymbol{u} \times \underset{\uparrow}{\boldsymbol{v}}) \\ &= \nabla \cdot (\underset{\uparrow}{\boldsymbol{u}} \times \boldsymbol{v}) - \nabla \cdot (\underset{\uparrow}{\boldsymbol{v}} \times \boldsymbol{u}),\end{aligned}$$

where the vertical arrow points to the argument on which ∇ operates. Since in the box product the dot and cross may be interchanged, we obtain

$$\nabla \cdot (\boldsymbol{u} \times \boldsymbol{v}) = (\nabla \times \boldsymbol{u}) \cdot \boldsymbol{v} - (\nabla \times \boldsymbol{v}) \cdot \boldsymbol{u}.$$

For the establishment of this formula by means of limits, see *Juvet* (11), I, p. 99.

As another example consider

$$\nabla \times (\boldsymbol{u} \times \boldsymbol{v}).$$

We first write

$$\nabla \times (\boldsymbol{u} \times \boldsymbol{v}) = \nabla \times (\underset{\uparrow}{\boldsymbol{u}} \times \boldsymbol{v}) + \nabla \times (\boldsymbol{u} \times \underset{\uparrow}{\boldsymbol{v}}).$$

Now apply the result of §**9.2** and obtain

$$\nabla \times (\underset{\uparrow}{\boldsymbol{u}} \times \boldsymbol{v}) = \begin{vmatrix} \boldsymbol{u} & \boldsymbol{v} \\ \nabla \cdot \boldsymbol{u} & \boldsymbol{v} \cdot \nabla \end{vmatrix}; \nabla \times (\boldsymbol{u} \times \underset{\uparrow}{\boldsymbol{v}}) = \begin{vmatrix} \boldsymbol{u} & \boldsymbol{v} \\ \boldsymbol{u} \cdot \nabla & \nabla \cdot \boldsymbol{v} \end{vmatrix}.$$

Hence

$$\nabla \times (\boldsymbol{u} \times \boldsymbol{v}) = (\boldsymbol{v} \cdot \nabla)\boldsymbol{u} - (\boldsymbol{u} \cdot \nabla)\boldsymbol{v} + \boldsymbol{u}(\nabla \cdot \boldsymbol{v}) - \boldsymbol{v}(\nabla \cdot \boldsymbol{u}).$$

Such formulas as these may of course be established by resorting to an $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ system. Cf. *Wills* (23), p. 93.

Exercises

17.6. Show that the operators ∇, $\nabla \cdot$, $\nabla \times$ are linear.

17.7. If $r = \sqrt{x^2 + y^2 + z^2}$, show that $1/r$ satisfies Laplace's equation.

17.8. If $V(P)$ is a vector field and $r(P)$ is the position vector of P, show that $(V \cdot \nabla)r = V$.

17.9. Show that the vector field $V(P)$, all of whose vectors are perpendicular to a fixed plane and whose magnitude is a function of the distance from the plane, is the gradient of a scalar field.

17.10. Show that the vector field all of whose vectors are perpendicular to a fixed line or axis and pass through that axis, and whose field intensity is a function of the distance from the axis, is the gradient of a scalar field.

17.11. Verify the following:
(1) div (grad f) $= \nabla^2 f$;
(2) div (curl V) $= 0$.

17.12. If r is the position vector of a point P referred to an i, j, k system, show that

$$\text{div } r = 3 \text{ and curl } r = \mathbf{0}.$$

17.13. By use of the identity

$$a \times (b \times c) = \begin{vmatrix} b & c \\ a \cdot b & a \cdot c \end{vmatrix},$$

or otherwise, show that

$$\nabla (u \cdot v) = v \times (\nabla \times u) + u \times (\nabla \times v) + (v \cdot \nabla)u + (u \cdot \nabla)v.$$

17.14. If $V(P)$ is a vector field and t is the unit tangent vector to the field lines, show that

$$t = \frac{V(P)}{|V|} \text{ and } \frac{1}{\rho} = |(t \cdot \nabla)t|,$$

$1/\rho$ being the curvature of the field line.

17.15. If V is the linear velocity vector of a point of a rigid body having an angular velocity vector w, show that

$$\text{curl } V = 2w \text{ and div } V = 0.$$

This interpretation makes the name "curl" seem appropriate. (Cf. *Gibbs-Wilson* (7), p. 155.) In place of "curl" the term "rotation" or "rotor" of the vector field is occasionally used, in which case it is written *rot V*.

§18. Divergence and Related Theorems

References: *Juvet* (11), Chapter V; *Wills* (23), Chapter IV. For a critical presentation of the material of this section, see *Kellogg* (40).

By use of the limit definitions of gradient, divergence, and curl, we are now in position to derive readily certain important theorems of the integral calculus which are particularly useful in physical and geometric applications.

18.1 Theorems of the gradient, divergence, and rotational.

Let T be a region of space in which the fields involved are defined and which is bounded by a closed surface Σ. We assume that the scalar field $f(P)$ has a gradient at each point of T. Let the region T be decomposed into N parts, $\Delta\tau_i$ being the volume of the ith part, which we supposed bounded by the surface Σ_i, and let P_i be an interior point of this subregion. Then by the definition of the gradient of $f(P)$ and the meaning of a limit, it follows that

$$\nabla f(P_i)\Delta\tau_i = \int_{\Sigma_i} d\mathbf{o} f(Q) + \boldsymbol{\varepsilon}_i\Delta\tau_i,$$

where $\boldsymbol{\varepsilon}_i$ is a vector which tends to zero as points Q on the boundary approach P. Let η be the greatest of the vectors $\boldsymbol{\varepsilon}_i$, that is, the one of greatest magnitude. Then η approaches zero as N approaches infinity, and each subregion approaches zero in all its dimensions. We now sum the above equations from one to N and pass to the limit. The result is

$$\int_T \nabla f(P)d\tau = \int_\Sigma d\mathbf{o} f(Q).$$

For

(1) $|\boldsymbol{\varepsilon}_1\Delta\tau_1 + \boldsymbol{\varepsilon}_2\Delta\tau_2 + \cdots + \boldsymbol{\varepsilon}_N\Delta\tau_N| \leqq |\eta|\{\Delta\tau_1 + \Delta\tau_2 + \cdots + \Delta\tau_N\} = |\eta|T$, which approaches zero as a limit, since T is finite;

(2) Over a "bulkhead" between adjoining subregions,

$$\int d\boldsymbol{\delta} f(Q) = \boldsymbol{0},$$

since the normals are oppositely pointed, and the integration is carried out twice over this surface.

The relation thus obtained

$$\int_T \nabla f(P) d\tau = \int_\Sigma d\boldsymbol{\delta} f(Q)$$

is called the *Theorem of the Gradient.*

Similar considerations yield the following two theorems:

$$\int_T \nabla \cdot \boldsymbol{V}(P) d\tau = \int_\Sigma d\boldsymbol{\delta} \cdot \boldsymbol{V}(Q)$$
$$\int_T \nabla \times \boldsymbol{V}(P) d\tau = \int_\Sigma d\boldsymbol{\delta} \times \boldsymbol{V}(Q).$$

The first of these is the extremely important relation known as the **Divergence Theorem**; we shall call the second the *Theorem of the Rotational.*

The divergence theorem may be stated: *The integral of the divergence of a field extended over a volume T bounded by a closed surface Σ is equal to the total flux of the field across Σ.*

18.2 Cartesian equivalent of the theorem of the gradient.

The Cartesian equivalents of these theorems are readily obtained. Let

$$\begin{aligned} f(P) &= f(x, y, z) \\ \boldsymbol{V}(P) &= X(x, y, z)\boldsymbol{i} + Y(x, y, z)\boldsymbol{j} + Z(x, y, z)\boldsymbol{k} \\ d\boldsymbol{\delta} &= d\sigma(\cos\alpha\, \boldsymbol{i} + \cos\beta\, \boldsymbol{j} + \cos\gamma\, \boldsymbol{k}), \end{aligned}$$

$d\sigma$ being an element of surface area.

Equating the coefficients of the vector $\boldsymbol{i}$ in the theorem of the gradient, we have

$$\int_T \frac{\partial f(P)}{\partial x} d\tau = \int_\Sigma \cos\alpha f(Q) d\sigma.$$

Now α is the angle between the outward-pointing normal $\boldsymbol{\xi}$ of the surface and $\boldsymbol{i}$, and hence $\cos\alpha\, d\sigma$ is the orthogonal

projection of the element of area $d\sigma$ of Σ on the y, z-plane. Since the integral is independent of the shape of the element of area, we may take $d\sigma$ such that its projection on the y, z-plane is the rectangle whose sides are dy and dz. Also we may take the element of volume to be $dx\,dy\,dz$. Therefore, the equivalent of the theorem of the gradient is

$$\int\int\int_T \frac{\partial f}{\partial x} dx\,dy\,dz = \int\int_\Sigma f\,dy\,dz,$$

with the two additional equations obtained by a cyclic permutation on the letters x, y, z.

Exercises

18.1. Show that the Cartesian form of the divergence theorem is

$$\int\int\int_T \left(\frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y} + \frac{\partial Z}{\partial z}\right) dx\,dy\,dz = \int\int_\Sigma (X dy\,dz + Y dz\,dx + Z dx\,dy).$$

Note.—This usually goes under the name of *Green's Theorem*, or the *Theorem of Ostrogradsky*. See books on analysis, such as *Osgood* (45); *Gibson* (33); *Goursat* (34); *Picard* (47), etc.

18.2. The Cartesian equivalent of the theorem of the rotational is

$$\int\int\int_T \left(\frac{\partial Z}{\partial y} - \frac{\partial Y}{\partial z}\right) dx\,dy\,dz = \int\int_\Sigma (Z dx\,dz - Y dy\,dx),$$

with two similar equations obtained by a cyclic permutation on the letters x, y, z.

18.3. Show that $\int_\Sigma d\mathbf{\sigma} = 0$, if Σ is a closed surface, by applying the theorem of the gradient.

18.4. Interpret the divergence theorem geometrically for

$$V = r = xi + yj + zk.$$

18.5. If $V = \text{grad } f$, prove by means of the divergence theorem the important theorem in the theory of harmonic functions

$$\int_T \nabla^2 f d\tau = \int_\Sigma d\boldsymbol{\sigma} \cdot \nabla f = \int_\Sigma \frac{df}{dn} d\sigma.$$

That is, the integral of the Laplacian of f extended over a volume T is equal to the flux of grad f across the bounding surface. Or, otherwise stated: The integral of the Laplacian of f over a volume T is equal to the integral over the bounding surface of the directional derivative of f in the direction of the outward-pointing normal to the surface.

18.6. Verify the divergence theorem for the case in which

$$V(P) = y\boldsymbol{i} - x^2\boldsymbol{j} + xz\boldsymbol{k},$$

and the region T is that bounded by the x, y-plane, $z > 0$; the y, z-plane, $x > 0$; and the unit sphere with center at the origin.

18.3 Stokes's theorem.

In each of the three theorems just considered an integral over a volume was expressed in terms of an associated integral over the closed surface enclosing that volume. We now obtain an important result which relates a certain surface integral to a line integral along a bounding curve.

We first consider a preliminary theorem. Let $\boldsymbol{a}$ be a constant unit vector, and consider the cylindrical surface which is generated by a line parallel to $\boldsymbol{a}$ as it traverses a

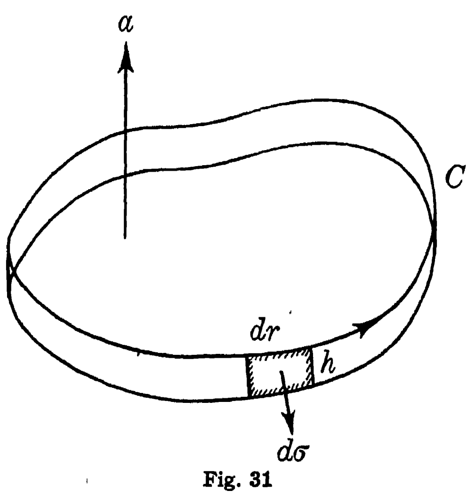

Fig. 31

simple closed curve C in a plane perpendicular to a (Fig. 31). Let Σ consist of the cylindrical surface between two planes each perpendicular to a and at a distance h apart together with these bases. From the definition of curl V,

$$\begin{aligned}
\mathbf{a} \cdot \text{curl } V &= \mathbf{a} \cdot (\nabla \times V) \\
&= \mathbf{a} \cdot \lim_{\text{points of } \Sigma \to P} \frac{\int_\Sigma d\boldsymbol{\sigma} \times V(Q)}{\tau} \\
&= \lim_{\text{points of } \Sigma \to P} \frac{\int_\Sigma \mathbf{a} \cdot (d\boldsymbol{\sigma} \times V(Q))}{\tau} \\
&= \lim_{\text{points of } \Sigma \to P} \frac{\int_\Sigma V \cdot (\mathbf{a} \times d\boldsymbol{\sigma}))}{\tau}.
\end{aligned}$$

Along the base of the cylindrical surface, $d\boldsymbol{\sigma}$ is parallel to $\boldsymbol{a}$ and hence $\boldsymbol{a} \times d\boldsymbol{\sigma} = \boldsymbol{0}$; on the curved surface $\boldsymbol{a} \times d\boldsymbol{\sigma} = h\,d\boldsymbol{r}$, where h is the height of the cylinder and $d\boldsymbol{r}$ is the vector line element of the bounding curve C. Hence

$$V \cdot (\mathbf{a} \times d\boldsymbol{\sigma}) = h\,V \cdot d\boldsymbol{r},$$

and therefore

$$\begin{aligned}
\mathbf{a} \cdot (\nabla \times V) &= \lim_{\text{points of } \Sigma \to P} \frac{h \int_C V \cdot d\boldsymbol{r}}{h\alpha} \\
&= \lim_{\text{points of } \Sigma \to P} \frac{\int_C V \cdot d\boldsymbol{r}}{\alpha},
\end{aligned}$$

α being the area of the base of the cylinder.

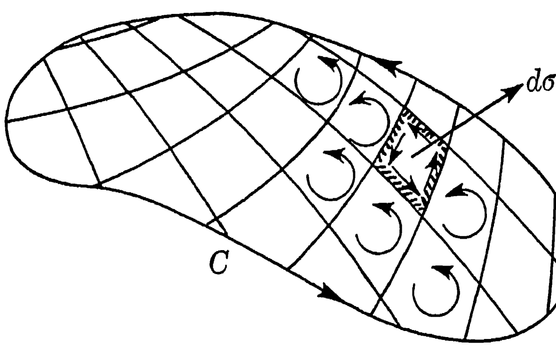

Fig. 32

Let S be a portion of a surface bounded by a simple closed curve C (Fig. 32). Let positive directions be assigned to the curve and the normal to the surface so that a rotation in the positive direction of C would advance a right-handed screw in the direction assigned to the surface normal $\boldsymbol{\xi}$. Let S be divided into subregions by a network of curves. For the ith subregion we have, by the above preliminary theorem,

$$\begin{aligned} d\boldsymbol{\sigma}_i \cdot (\nabla \times V(P_i)) &= \boldsymbol{\xi}_i \cdot (\nabla \times V(P_i)) d\sigma_i \\ &= \int_{C_i} V(Q) \cdot d\boldsymbol{r} + \epsilon \Delta\sigma_i, \end{aligned}$$

where ϵ approaches zero as $\Delta\sigma_i$ approaches zero. Summing and passing to the limit as each subregion shrinks toward an interior point, we have

$$\int_S d\boldsymbol{\sigma} \cdot (\nabla \times V) = \int_C d\boldsymbol{r} \cdot V,$$

which is known as the formula of *Ampere-Stokes*, or **Stokes's Theorem.** In words, *the total flux of the curl of a vector field V across a surface S bounded by a curve C is equal to the total circulation of V along C.*[1]

Exercises

18.7. Obtain Green's theorem in the plane

$$\int\int_S \left(\frac{\partial Y}{\partial x} - \frac{\partial X}{\partial y}\right) dx\, dy = \int_C (X dx + Y dy)$$

as an application of Stokes's theorem.

18.8. Establish the results

$$\int_S d\boldsymbol{\sigma} \times \nabla f = \int_C d\boldsymbol{r} f$$

$$\int_S (d\boldsymbol{\sigma} \times \nabla) \times V = \int_C d\boldsymbol{r} \times V.$$

Hint.—Apply Stokes's theorem to $f\,\mathbf{a}$ and to $V \times \mathbf{a}$, where $\mathbf{a}$ is an *arbitrary* but *constant* vector field.

[1] *Kampen, E. R. van,* "The Theorems of Gauss-Bonnet and Stokes," *American Journal of Mathematics*, *60*, pp. 129–138 (1938).

18.9. If $\boldsymbol{r}$ is the position vector of points P on a closed curve C, show that

$$\int_C \boldsymbol{r} \cdot d\boldsymbol{r} = 0.$$

18.10. Interpret geometrically the integral

$$\int_C d\boldsymbol{r} \times \boldsymbol{r},$$

where $\boldsymbol{r}$ is the position vector of points on a curve C which is plane and closed.

18.11. If f and g are scalar fields defined over a surface S bounded by a closed curve C, show that

$$\int_C f \operatorname{grad} g \cdot d\boldsymbol{r} = -\int_C g \operatorname{grad} f \cdot d\boldsymbol{r}.$$

18.12. If Σ is a closed surface, show that

$$\int_\Sigma d\boldsymbol{\sigma} \cdot \operatorname{curl} \boldsymbol{V} = 0, \text{ and } \int_\Sigma d\boldsymbol{\sigma} \times \operatorname{grad} f = \boldsymbol{0}.$$

18.13. As a special instance of Green's theorem in the plane, show that the plane area bounded by a simple closed curve C is given by the line integral

$$\tfrac{1}{2}\int_C (x\,dy - y\,dx).$$

Cf. *Goursat* (34), p. 187.

18.14. Establish the following relations, each of which is known as Green's theorem.

If $\boldsymbol{V} = f \operatorname{grad} g$, then $\operatorname{div} \boldsymbol{V} = f\,\nabla^2 g + \operatorname{grad} f \cdot \operatorname{grad} g$, and $\boldsymbol{V} \cdot d\boldsymbol{\sigma} = f \operatorname{grad} g \cdot d\boldsymbol{\sigma} = f\,dg/dn$.

The theorems to be established are

$$(1) \qquad \int_T (f\nabla^2 g + \operatorname{grad} f \cdot \operatorname{grad} g)\,d\tau = \int_\Sigma f\frac{dg}{dn}\,d\sigma$$

$$(2) \qquad \int_T (f\nabla^2 g - g\nabla^2 f)\,d\tau = \int_\Sigma \left(f\frac{dg}{dn} - g\frac{df}{dn}\right)d\sigma,$$

where Σ is the closed surface bounding the region T.

18.15. Prove that if there is applied to each element ds of a closed and rigid curve C a force of magnitude ds/ρ in the direction

of the principal normal vector, the curve remains in equilibrium; $1/\rho$ denotes the curvature of the curve and s is its arc length.

Hint.—The translational effect of the applied forces is given by $\int_C \frac{\xi_2}{\rho} ds$ and the rotational effect by $\int_C \left(r \times \frac{\xi_2}{\rho}\right) ds$, where ξ_2 is the unit principal normal vector of the curve and r is the position vector of a point of C. Hence the vanishing of these integrals is equivalent to the curve remaining in equilibrium.

18.16. A rigid body of volume T bounded by the surface Σ is completely immersed in a fluid of specific gravity unity. Prove that the effect of the fluid pressure on the body is the same as that of a single force f of magnitude T, vertically upwards, applied at the centroid C of the volume T.

Hint.—Let the x, y-plane be the surface of the liquid. To obtain the resultant force vector apply the Theorem of the Gradient to $\int_\Sigma z\, d\boldsymbol{\sigma}$, and to obtain its line of application consider the moment with respect to the origin and apply the Theorem of the Rotational.

§19. Examples of Applications

19.1 Line integral independent of the path.

Let C be a curve joining two fixed points A, B and consider the integral $\int_C d\boldsymbol{r} \cdot \boldsymbol{V}$, where V is a vector field. A necessary and sufficient condition that this integral shall have a value which depends only on the end points A and B, but is independent of the curve joining them, is that the integral taken along an arbitrary closed curve passing through A and B shall have the value zero. Suppose then

$$\int_{C^*} d\boldsymbol{r} \cdot \boldsymbol{V} = 0,$$

where C^* is an arbitrary closed curve passing through A and B. By Stokes's theorem it follows that

$$\int_\Sigma d\boldsymbol{\sigma} \cdot (\nabla \times \boldsymbol{V}) = 0,$$

where Σ is any surface bounded by the closed curve C^*. Obviously a sufficient condition that this equation be true

is that

$$\nabla \times V = 0.$$

On account of the *arbitrariness* of the surface Σ, we conclude that it is also necessary.

An equivalent statement of this theorem is (see exercises below): *A necessary and sufficient condition that a vector field admit of being the gradient of a scalar field is that its curl shall vanish.* The function $f(x, y, z)$ defining the scalar field, or frequently its negative, is called a *potential function* for the vector field which is its gradient. A vector field admitting a potential function is called a *conservative field.*

As an example, consider the field of force acting on a unit mass due to a mass particle of mass M under the Newtonian law of gravitation. Let the mass M be situated at O, and denote $\overrightarrow{OP}$ by $\mathbf{r}$. Then

$$V = -\frac{\lambda M \mathbf{r}}{r^3},$$

where λ is a constant.

Now

$$d\mathbf{r} \cdot V = -\frac{\lambda M(\mathbf{r} \cdot d\mathbf{r})}{r^3} = -\frac{\lambda M\, dr}{r^2} = d\left(\frac{\lambda M}{r}\right).$$

Hence $\lambda M/r$ is a potential function for V, since

$$d\frac{\lambda M}{r} = d\mathbf{r} \cdot \operatorname{grad}\frac{\lambda M}{r} = d\mathbf{r} \cdot V,$$

for $d\mathbf{r}$ arbitrary. We write

$$U(P) = \frac{\lambda M}{r},$$

which is defined, together with its derivatives, at all points P other than O. Thus the Newtonian field is conservative and the work done in transporting a unit particle from a point A to a point B is independent of the path.

The total flux of V across a closed surface Σ is

$$\int_\Sigma d\boldsymbol{\delta} \cdot V = -\lambda M \int_\Sigma \frac{d\boldsymbol{\delta} \cdot \boldsymbol{r}}{r^3}.$$

We recognize that this is just the integral which gives the solid angle of Σ when viewed from O. Hence the flux of V is 0 or $-4\pi\lambda M$ according as O is an exterior or interior point of Σ.

Let P be a point distinct from O, and let P be surrounded by a closed surface Σ which does not contain O as an interior point. The divergence of V at P is defined by

$$\lim_{\text{points of } \Sigma \to P} \frac{\int_\Sigma d\boldsymbol{\delta} \cdot V(Q)}{\tau}.$$

Since the numerator is zero, it follows that div $V = 0$ at all points P distinct from O. Hence at points of free space, the Newtonian potential function due to a point mass satisfies Laplace's equation

$$\nabla^2 U = 0.$$

These considerations show that the same is true for the Newtonian potential arising from any finite number of discreet point masses. In the case of continuous masses, the vector field V is defined by the integral

$$V(P) = -\int_T \frac{\rho(Q)\boldsymbol{r}}{r^3} d\tau$$

extended over the volume T, where $\rho(Q)$ is a density factor and where we have taken λ to be equal to 1. The potential function $U(P)$ is found to be

$$U(P) = \int_T \frac{\rho(Q)}{r} d\tau,$$

and it can be shown (*Juvet* (12), p. 20) that

$$V(P) = \nabla U(P)$$

at *every point of space.*

Exercises

19.1. A necessary and sufficient condition that $\boldsymbol{V} \cdot d\boldsymbol{r}$ be an exact differential is that $\boldsymbol{V}$ be the gradient of a scalar point function.

19.2. A necessary and sufficient condition that $\boldsymbol{V}$ be the gradient of a scalar is that curl $\boldsymbol{V} = \boldsymbol{0}$.

19.3. If a potential function exists for a given vector field, it is unique except for an additive constant.

19.4. If $\boldsymbol{V} \cdot d\boldsymbol{r}$ is not an exact differential, a necessary condition that there exist a scalar $\varphi(x, y, z)$ such that $\varphi\boldsymbol{V} \cdot d\boldsymbol{r}$ is an exact differential is that the field $\boldsymbol{V}$ be perpendicular to its curl at each point. (The condition is also sufficient.) The theorem may also be stated: Given a vector field, a necessary (and sufficient) condition that there exist a family of surfaces orthogonal to the field lines is that the vector field be perpendicular to its curl at each point.

19.2 Physical interpretation of divergence.

References: *Wills* (23), p. 106; *Webster* (52), p. 496; *Kellogg* (40), p. 45; *Gibbs-Wilson* (7), p. 152.

Let $\boldsymbol{V}$ be the velocity of a moving fluid and let ρ be its density, both of which we shall regard as functions of position x, y, z and the time t. Let Σ be a closed surface, fixed in space, which bounds a region T through which the fluid is flowing. The total mass of the fluid in the region T at a time t is given by

$$M(t) = \int_T \rho\,(t, x, y, z)d\tau,$$

and its rate of change per unit time is given by

$$\frac{dM(t)}{dt} = \int_T \frac{\partial\rho}{\partial t}d\tau.$$

But the rate at which fluid is entering the region T is given by the surface integral

$$-\int_\Sigma d\boldsymbol{\sigma} \cdot \rho\boldsymbol{V},$$

where $d\boldsymbol{\sigma}$ has the direction of the outward-pointing normal. Hence

$$-\int_{\Sigma} d\boldsymbol{\sigma} \cdot \rho V = \int_T \frac{\partial \rho}{\partial t} d\tau.$$

Applying the divergence theorem to the integral on the left, we have

$$-\int_T \operatorname{div} (\rho V) d\tau = \int_T \frac{\partial \rho}{\partial t} d\tau,$$

or

$$\int_T \left\{ \frac{\partial \rho}{\partial t} + \operatorname{div} (\rho V) \right\} d\tau = 0.$$

Since this relation holds for an arbitrary region T, we conclude that

$$\frac{\partial \rho}{\partial t} + \operatorname{div} (\rho V) = 0$$

at each point of the fluid and for an arbitrary time t. This is known as the *equation of continuity* of a perfect fluid. Expanding,

$$\frac{\partial \rho}{\partial t} + (\nabla \rho) \cdot V + \rho(\nabla \cdot V) = 0,$$

or

$$\frac{\partial \rho}{\partial t} + \frac{\partial \rho}{\partial x} \frac{dx}{dt} + \frac{\partial \rho}{\partial y} \frac{dy}{dt} + \frac{\partial \rho}{\partial z} \frac{dz}{dt} + \rho \operatorname{div} V = 0,$$

or

$$\frac{\dfrac{d\rho}{dt}}{\rho} + \operatorname{div} V = 0.$$

Consider now a small portion of the fluid of volume τ *as it moves;* its mass will be a constant, say $m = \rho\tau$. Taking the logarithm of each side and differentiating with respect to t, we have

$$0 = \frac{\dfrac{d\rho}{dt}}{\rho} + \frac{\dfrac{d\tau}{dt}}{\tau}.$$

Hence, from the equation of continuity

$$\operatorname{div} V = \frac{\frac{d\tau}{dt}}{\tau},$$

that is, the *divergence of the velocity is the time rate of increase of volume per unit volume.*.

If $d\rho/dt = 0$, the fluid is said to be *incompressible.* Hence for an incompressible fluid

$$\operatorname{div} V = 0.$$

If ρ is constant not only with respect to time but also with respect to position, the fluid is of uniform density and the

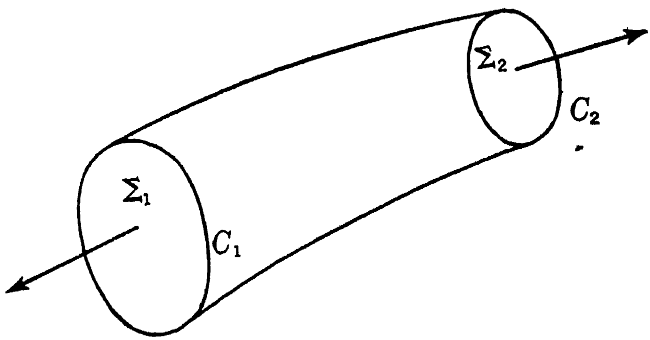

Fig. 33

mass of any portion of it is proportional to the volume. In this case

$$\int_{\Sigma} d\sigma \cdot V = 0$$

expresses the fact that the volume of the fluid entering a given region per unit time is equal to the volume leaving it.

Let V be a vector field whose divergence vanishes in a region R. Let C be a closed curve in R. Then the field lines of V passing through points of C will form a *tube.* Let Σ_1 and Σ_2 be two surfaces cutting the tube. We have then a closed surface S consisting of the curved surface of the tube and its two end surfaces Σ_1 and Σ_2. At each point on the curved surface of the tube the normal to the surface

is perpendicular to the field line at that point. Then applying the divergence theorem

$$\int_{\Sigma_1} d\mathfrak{d} \cdot V + \int_{\Sigma_2} d\mathfrak{d} \cdot V = 0.$$

If we take the *inward*-pointing normal to Σ_1, we have

$$\int_{\Sigma_1} d\mathfrak{d} \cdot V = \int_{\Sigma_2} d\mathfrak{d} \cdot V,$$

which says that the flux of V across a section of the tube is constant. This is realized, for instance, in the flow of an incompressible liquid of constant density through a pipe.

Exercise

19.5. A necessary condition that the differential equation

$$\nabla \times V = a$$

where a is a given vector field, admit a solution V is that div $a = 0$. If V is a solution, then $V +$ grad f is also a solution where f is an arbitrary scalar point function.

19.3 On the flow of heat.

Reference: *Kellogg* (40), p. 76 *ff.*

Suppose a solid all of whose points are not at the same temperature. The rate of flow of heat may be represented by a vector field V whose direction at any point P is that in which heat is flowing, and whose magnitude is obtained by taking an element $\Delta\sigma$ of the plane through the point P normal to the direction of flow, determining the number of calories per second flowing through this element, dividing this number by the area $\Delta\sigma$, and taking the limit of the quotient as points Q of $\Delta\sigma$ approach P. We assume: (1) that the velocity of flow is proportional to the rate of fall of the temperature U at P, where the proportionality factor depends on the conductivity of the material; (2) that the body is *thermally isotropic.* Hence the flow vector V has the same direction as the gradient of U and the opposite sense,

$$V = -\lambda \text{ grad } U.$$

That is, we have made the assumption that the flow of heat is orthogonal to the isothermal surfaces.

Consider now any simply connected closed region T in the body, and compare the rate of flow of heat into T against the rise in temperature. The rate of flow of heat into T in calories per unit time is given by

$$-\int_{\Sigma} d\delta \cdot V,$$

where Σ is the closed surface bounding T. A calorie of heat will raise a unit mass of the body c degrees, c being the specific heat of the material. Hence the number of calories per unit time received per unit of mass is measured by

$$c\frac{\partial U}{\partial t},$$

and the number of calories received in unit time by the whole mass in T is

$$\int_T c\rho \frac{\partial U}{\partial t} d\tau,$$

where ρ is the density function for the material. Hence

$$\int_T c\rho \frac{\partial U}{\partial t} d\tau + \int_{\Sigma} d\delta \cdot V = 0,$$

or, by means of the divergence theorem,

$$\int_T \left\{ c\rho \frac{\partial U}{\partial t} + \operatorname{div} V \right\} d\tau = 0.$$

Since this equation holds for an arbitrary region, we conclude that

$$\frac{\partial U}{\partial t} = -\frac{1}{c\rho} \operatorname{div} V,$$

or since

$$V = -\lambda \operatorname{grad} U$$

we have

$$\mu^2 \nabla^2 U = \frac{\partial U}{\partial t},$$

where μ is a constant on the assumption that λ, c, and ρ are constants.

If a stationary state of temperature has been established, $\partial U/\partial t = 0$, and the temperature function U then satisfies Laplace's equation

$$\nabla^2 U = 0.$$

CHAPTER IV

Introduction to Tensor Analysis

§20. Tensors and Invariants

The literature relating to tensor analysis is very extensive. From the great wealth of material the following references are especially recommended in this connection. *Veblen* (50); *Juvet* (38); *Eddington* (31); *Murnaghan* (44); *Cartan* (28).

20.1 Coördinate system and N-dimensional space.

Reference: *Veblen* (50), p. 13.

By a *space of N-dimensions* we shall mean a set of objects, usually called *points*, which is in a one-to-one reciprocal correspondence with the totality of ordered sets of N real numbers $(x^1, x^2, \ldots, x^N)$ satisfying a set of inequalities

$$|x^\alpha - A^\alpha| < k^\alpha, \ (\alpha = 1, 2, \cdots, N)$$

where the A^α are constants and the k^α are positive constants. The correspondence is called a *coördinate system.* The numbers $x^1, x^2, \ldots, x^N$ are called the *coördinates* of the point to which the set $(x^1, x^2, \ldots, x^N)$ corresponds in the coördinate system. From a notational standpoint it proves to be more convenient to write the coördinates with superscripts, as we have done, rather than with subscripts.

20.2 Transformation of coördinates.

References: *Veblen* (50), p. 13; *Goursat* (34), pp. 399–407 and Chapter II.

A set of N equations

$$y^i = y^i(x^1, x^2, \cdots, x^N), \ (i = 1, 2, \cdots, N)$$

in which the y^i are single-valued functions for all points $(x^1, x^2, \ldots, x^N)$ in a region R of the N-dimensional space, and which admit of a unique solution for the x's in terms of the y's

$$x^\alpha = x^\alpha(y^1, y^2, \cdots, y^N), \quad (\alpha = 1, 2, \cdots, N),$$

which are also single-valued functions over the same region R will be called a *transformation of coördinates*. A transformation of coördinates is a device which relabels each point with new coördinates in an unambiguous fashion. Thus, if the coördinates of a point with respect to one coördinate system are $(x^1, x^2, \ldots, x^N)$, the numbers y^1, $y^2, \ldots, y^N$ computed from these equations are the coördinates of the *same* point with respect to the other coördinate system.

For simplicity, we shall require the admitted transformation to be *analytic* in the region R under consideration. We also suppose the functional determinant

$$\left|\frac{\partial y}{\partial x}\right| \equiv \begin{vmatrix} \frac{\partial y^1}{\partial x^1} & \frac{\partial y^1}{\partial x^2} & \cdots & \frac{\partial y^1}{\partial x^N} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial y^N}{\partial x^1} & \frac{\partial y^N}{\partial x^2} & \cdots & \frac{\partial y^N}{\partial x^N} \end{vmatrix}$$

to be different from zero at each point of the region R. It can be proved that the class of transformations with these properties constitutes a *group* with respect to the operation of forming their resultant.

20.3 Invariants.

Any object which maintains its identity under the group of transformations is called an *invariant* with respect to the group of transformations. An example of an invariant is a point; its description changes, that is, its coördinates change under the transformation, but we can identify the *same* point after the transformation. A scalar point function is an invariant which is specified by a single quantity in

a given coördinate system, say $\varphi(x)$, where we write $\varphi(x)$ for $\varphi(x^1, x^2, \ldots, x^N)$. Its specification in any other coördinate system is given by

$$\bar{\varphi}(y) = \varphi[x^1(y^1, y^2, \cdots, y^N), x^2(y^1, y^2, \cdots, y^N), \cdots, x^N(y^1, y^2, \cdots, y^N)].$$

20.4 Definition of a vector.

Let x be a coördinate system, and let y be *any* coördinate system obtainable from the x-coördinate system by an admitted transformation. By a *contravariant vector* is meant an *invariant* (with respect to the admitted group of transformations) which has for its specification N components in each coördinate system $\xi^1(x), \xi^2(x), \ldots, \xi^N(x)$ and $\bar{\xi}^1(y), \bar{\xi}^2(y), \ldots, \bar{\xi}^N(y)$, which are related thus:

$$\bar{\xi}^i(y) = \xi^1(x)\frac{\partial y^i}{\partial x^1} + \xi^2(x)\frac{\partial y^i}{\partial x^2} + \cdots + \xi^N(x)\frac{\partial y^i}{\partial x^N}, \quad (i = 1, 2, \cdots, N).$$

We write this in the form

$$(A) \qquad \bar{\xi}^i(y) = \xi^\alpha(x)\frac{\partial y^i}{\partial x^\alpha}, \quad (i = 1, 2, \cdots, N),$$

with the understanding that the *repeated* index α is to be summed from 1 to N. In the following we shall adhere to this convention, which is now well standardized in tensor analysis. An index which is summed, as in this instance, is sometimes called a "dummy" index or "umbral" index. Such an index may be changed at will without effecting the result; it is in this respect like the variable of integration in a definite integral. Thus

$$\xi^\alpha\frac{\partial y^i}{\partial x^\alpha} = \xi^\beta\frac{\partial y^i}{\partial x^\beta} = \xi^\gamma\frac{\partial y^i}{\partial x^\gamma}, \text{ etc.}$$

In the equation (A) from a notation standpoint it is essential that the free index i occur on each side of the equation in the same relative position with respect to "upper" and "lower" index positions.

By a *covariant vector* is meant an *invariant* which has for its specification N components in each coördinate system which are related thus

$$(B) \qquad \bar{\eta}_i(y) = \eta_\alpha(x)\frac{\partial x^\alpha}{\partial y^i}, \quad (i = 1, 2, \cdots, N).$$

In practice the components of a vector are usually functions of the coördinate variables and their differentials. The components of a vector of either type can clearly be arbitrarily assigned in one coördinate system, but then the components of the *same* vector in any other admitted coördinate system are uniquely determined by these prescribed laws of transformation.

From the differential calculus

$$dy^i = dx^\alpha \frac{\partial y^i}{\partial x^\alpha},$$

where of course α is summed from 1 to N and where i takes on the set of values $1, 2, \ldots, N$. (In the future all indices will be understood to have the range $1, 2, \ldots, N$ unless otherwise indicated.) From these equations we see that the differentials of the coördinates $dx^1, dx^2, \ldots, dx^N$ constitute a *contravariant vector.* This may be regarded as the "typical" contravariant vector. As a memory device for (A), we need only recall that the components of any contravariant vector transform exactly the same way as the differentials of the coördinates.

Let $\varphi(x)$ be a scalar point function. Then by the differential calculus,

$$\frac{\partial \varphi}{\partial y^i} = \frac{\partial \varphi}{\partial x^\alpha}\frac{\partial x^\alpha}{\partial y^i}.$$

Hence the quantities

$$\frac{\partial \varphi}{\partial x^1}, \frac{\partial \varphi}{\partial x^2}, \cdots, \frac{\partial \varphi}{\partial x^N}$$

constitute a *covariant vector.* This covariant vector is called the *gradient* of the scalar field. It may be regarded

as the typical covariant vector, and this fact used as a memory device for the transformation (B).

In case a set of N quantities constitute a vector, we shall always use upper indices (superscripts) in the case of a contravariant vector, and lower indices (subscripts) in the case of a covariant vector. Thus the notation distinguishes the two types of vectors, that is, the two types of description (see §9.5).

Exercises

20.1. The relation between rectangular Cartesian and polar coördinates in the plane is given by

$$x = r \cos \theta$$
$$y = r \sin \theta.$$

Determine a region of the plane for which these equations define an admitted transformation of coördinates.

20.2. If x and y are coördinates related by a transformation, establish the following identities:

$$\frac{\partial y^i}{\partial x^\alpha} \frac{\partial x^\alpha}{\partial y^j} = \delta^i_j, \tag{1}$$

where the symbol δ^i_j, known as the *Kronecker delta*, is defined by

$$\delta^i_j = \begin{cases} 1 \text{ for } i = j. \\ 0 \text{ for } i \neq j. \end{cases}$$

$$\frac{\partial^2 y^i}{\partial x^\alpha \partial x^\beta} \frac{\partial x^\alpha}{\partial y^j} + \frac{\partial y^i}{\partial x^\alpha} \frac{\partial^2 x^\alpha}{\partial y^j \partial y^k} \frac{\partial y^k}{\partial x^\beta} = 0. \tag{2}$$

$$\frac{\partial^2 y^i}{\partial x^\alpha \partial x^\beta} \frac{\partial x^\beta}{\partial y^k} \frac{\partial x^\alpha}{\partial y^j} + \frac{\partial y^i}{\partial x^\alpha} \frac{\partial^2 x^\alpha}{\partial y^j \partial y^k} = 0. \tag{3}$$

20.3. The velocity vector dx^α/dt is a contravariant vector; however, the quantities d^2x^α/dt^2 do *not* constitute a vector, and hence are not the components of the acceleration vector.

20.4. If x is an affine coördinate system determine the most general group of transformation of coördinates under which d^2x^α/dt^2 is a contravariant vector.

20.5. If ξ^α is a contravariant vector and η_β is a covariant vector, then $\xi^\alpha \eta_\alpha$ is a *scalar invariant.* This is called the "scalar product" of the two vectors.

20.6. If φ is a scalar point function, then $d\varphi$ is a scalar invariant.

20.7. With respect to the group of orthogonal transformations with determinant $+1$, the expressions for grad f, div V, and curl V of **§17** are invariant. (The dimensionality $N = 3$ is implied.)

20.8. The laws of transformation (A) and (B) may be equally well written in the equivalent forms

$$(A') \qquad \bar{\xi}^i \frac{\partial x^\alpha}{\partial y^i} = \xi^\alpha$$

$$(B') \qquad \bar{\eta}_i \frac{\partial y^i}{\partial x^\alpha} = \eta_\alpha.$$

20.9. If x, y are rectangular Cartesian coördinates and r, θ are polar coördinates in the plane, obtain the components in the polar coördinates of the vectors which are described in the x, y-coördinates by

$$\xi^1(x) = x^2 y, \qquad \xi^2(x) = x - y^2$$

and

$$\eta_1(x) = x^2 y, \qquad \eta_2(x) = x - y^2.$$

Obtain the particular vectors at the point $x = 1$, $y = 2$. Verify in this instance that $\xi^1\eta_1 + \xi^2\eta_2$ is a scalar invariant.

20.5 Tensors.

An *invariant* which is specified by N^2 components in each coördinate system $\xi^{\alpha\beta}(x)$ and $\bar{\xi}^{ij}(y)$ which transform according to

$$(C) \qquad \bar{\xi}^{ij}(y) = \xi^{\alpha\beta}(x) \frac{\partial y^i}{\partial x^\alpha} \frac{\partial y^j}{\partial x^\beta},$$

where i, j range independently from 1 to N, is called a *contravariant tensor of the second order* (or rank).

A *covariant tensor* of the second order is an invariant whose N^2 components transform according to

$$(D) \qquad \bar{\xi}_{ij}(y) = \xi_{\alpha\beta}(x) \frac{\partial x^\alpha}{\partial y^i} \frac{\partial x^\beta}{\partial y^j}.$$

A *mixed tensor* of the second order with one contravariant index and one covariant index is an invariant whose com-

ponents satisfy the law of transformation

$$(E) \qquad \xi^{i}{}_{j}(y) = \xi^{\alpha}{}_{\beta}(x)\frac{\partial y^i}{\partial x^\alpha}\frac{\partial x^\beta}{\partial y^j}.$$

Since the indices are ordered, we use the notation $\xi^{\alpha}{}_{\beta}$ and not ξ^{α}_{β}. The reason for this distinction will be more apparent when we consider the operation of raising and lowering indices.

The extension of the definitions to a tensor of any order is immediate. For example, a tensor of the fifth order with three contravariant indices and two covariant indices is an invariant whose N^5 components satisfy the law of transformation

$$\xi^{ij}{}_{k}{}^{l}{}_{m} = \xi^{\alpha\beta}{}_{\gamma}{}^{\lambda}{}_{\mu}\frac{\partial y^i}{\partial x^\alpha}\frac{\partial y^j}{\partial x^\beta}\frac{\partial x^\gamma}{\partial y^k}\frac{\partial y^l}{\partial x^\lambda}\frac{\partial x^\mu}{\partial y^m}.$$

A vector is a tensor of rank 1; a scalar may be regarded as a tensor of rank zero.

In tensor analysis we have an example of a marvelous notation. As stated above, the indices which are not summed are known as *free* indices. The same free indices must appear on each side of a tensor equation in the same relative positions. Thus we have a constant check similar to the employment of "dimensions" in certain physical theories. The notation is so condensed that an effort is required to appreciate the number of terms that may be involved in the description of a tensor. For instance, in the relativity theory $N = 4$ the curvature tensor carries four indices and hence has $N^4 = 256$ components. Two tensors of the same type are said to be equal if and only if their corresponding components are equal when the tensors are expressed in the same coördinate system.

Exercises

20.10. The equations

$$\bar{\xi}^{ij} = \xi^{\alpha\beta}\frac{\partial y^i}{\partial x^\alpha}\frac{\partial y^j}{\partial x^\beta}, \quad \bar{\xi}^{ij}\frac{\partial x^\beta}{\partial y^j} = \xi^{\alpha\beta}\frac{\partial y^i}{\partial x^\alpha}, \quad \bar{\xi}^{ij}\frac{\partial x^\alpha}{\partial y^i}\frac{\partial x^\beta}{\partial y^j} = \xi^{\alpha\beta},$$

are equivalent.

20.11. If ξ^α and η^β are contravariant vectors, the quantities $\zeta^{\alpha\beta}$ defined by $\zeta^{\alpha\beta} = \xi^\alpha\eta^\beta$ constitute a contravariant tensor of the second order.

20.12. If ξ^α, η^β, and $\varphi_{\lambda\mu}$ are tensors of the types indicated by the indices, then $\varphi_{\alpha\beta}\xi^\alpha\eta^\beta$ is a scalar invariant.

20.13. If $\varphi_{\alpha\beta} = \varphi_{\beta\alpha}$ for every pair of indices α, β the tensor $\varphi_{\alpha\beta}$ is said to be *symmetric.* Show that the property of $\varphi_{\alpha\beta}$ being symmetric is invariant with respect to the group of transformations.

20.6 Introduction of a metric.

References: *Weyl* (53), Chapter II; *Murnaghan* (44), Chapter III.

Let $g_{\alpha\beta}(x)$ be a symmetric covariant tensor of the second order whose components are functions of the coördinates only but do not involve their differentials. Then, as we have seen,

$$g_{\alpha\beta}dx^\alpha dx^\beta$$

is a scalar invariant. We suppose that $g_{\alpha\beta}$ are such that $g_{\alpha\beta}dx^\alpha dx^\beta$ is a *positive definite* quadratic differential form.

Let C be a curve defined by

$$C: \quad x^\alpha = x^\alpha(t).$$

Then the integral along C

$$s(t) = \int_{t_0}^{t} \sqrt{g_{\alpha\beta}\frac{dx^\alpha}{dt}\frac{dx^\beta}{dt}}\,dt$$

is a scalar invariant; that is, it has a value independent of the particular coördinate system used. We take the value of this integral, by definition, as the *length of arc of the curve* C from the point specified by t_0 to that given by t. The differential of arc is then given by

$$ds^2 = g_{\alpha\beta}dx^\alpha dx^\beta.$$

The tensor $g_{\alpha\beta}$ is called the *fundamental covariant tensor* with respect to the metrical properties of the space, for

it is what determines them according to the following definitions:

If ξ^α is a contravariant vector, its *length* (or magnitude) is defined by

$$\text{length of } \xi = +\sqrt{g_{\alpha\beta}\xi^\alpha\xi^\beta}.$$

If ξ^α and η^β are two contravariant vectors, the angle θ between them is defined by

$$\cos\theta = \frac{g_{\alpha\beta}\xi^\alpha\eta^\beta}{\sqrt{g_{\alpha\beta}\xi^\alpha\xi^\beta}\sqrt{g_{\alpha\beta}\eta^\alpha\eta^\beta}}.$$

The fundamental covariant tensor $g_{\alpha\beta}$ may be given arbitrarily, except for the conditions imposed at the

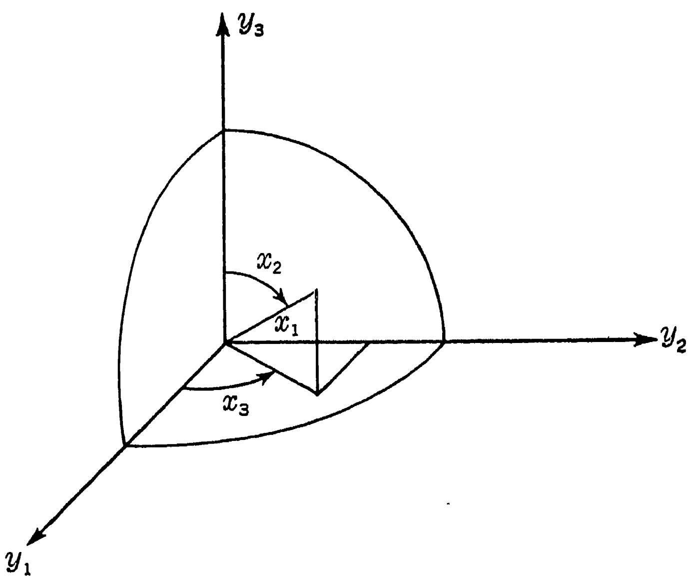

Fig. 34

beginning of this section, in one coördinate system; but, being a tensor, it is then completely determined in any related coördinate system. We shall illustrate this by

computing the fundamental tensor for the space spherical coördinate system. (Fig. 34.)

Let y^1, y^2, y^3 be rectangular Cartesian coördinates in a Euclidean space of three dimensions with the fundamental quadratic differential form

$$ds^2 = (dy^1)^2 + (dy^2)^2 + (dy^3)^2.$$

The fundamental tensor $\bar{g}_{ij}$ then has the components

$$\bar{g}_{11} = \bar{g}_{22} = \bar{g}_{33} = 1; \; \bar{g}_{12} = \bar{g}_{23} = \bar{g}_{31} = 0.$$

Let x^1, x^2, x^3 be space polar coördinates as indicated in the figure. The relations between the coördinate systems are

$$\begin{cases} y^1 = x^1 \sin x^2 \cos x^3 \\ y^2 = x^1 \sin x^2 \sin x^3 \\ y^3 = x^1 \cos x^2 \end{cases} \quad \text{and} \quad \begin{cases} x^1 = +\sqrt{(y^1)^2 + (y^2)^2 + (y^3)^2} \\ x^2 = \operatorname{arc\,tan} \dfrac{\sqrt{(y^1)^2 + (y^2)^2}}{y^3} \\ x^3 = \operatorname{arc\,tan} \dfrac{y^2}{y^1} \end{cases}$$

which constitute an allowable transformation of coördinates in the region of space satisfying the inequalities

$$x^1 > 0; \; 0 < x^2 < \pi; \; 0 \leqq x^3 < 2\pi.$$

From the law of transformation

$$g_{\alpha\beta} = \bar{g}_{ij} \frac{\partial y^i}{\partial x^\alpha} \frac{\partial y^j}{\partial x^\beta}$$

we obtain by straightforward computation

$$g_{11} = 1, \; g_{22} = (x^1)^2, \; g_{33} = (x^1 \sin x^2)^2,$$
$$g_{12} = g_{23} = g_{31} = 0.$$

Hence the fundamental quadratic differential form in the space polar coördinates is

$$ds^2 = (dx^1)^2 + (x^1)^2(dx^2)^2 + (x^1 \sin x^2)^2(dx^3)^2.$$

20.7 Local system of base vectors.

Through each point M of the N-dimensional space there pass N parametric lines. By a parametric line is meant a curve along which only one coördinate varies; if x^α is the

variable coördinate, the curve is called the "x^α parametric line." At an arbitrary point M, we set up a system of base vectors $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_N$ such that $\boldsymbol{e}_\alpha$ is tangent to the x^α parametric line. Consider an infinitesimal displacement in the space from the point M along the x^α parametric line. Such a displacement is specified by

$$dx^\beta = 0 \text{ for } \beta \neq \alpha,\ dx^\alpha \neq 0.$$

According to the definition of $\boldsymbol{e}_\alpha$, the vector representing the displacement is a scalar multiple of $\boldsymbol{e}_\alpha$. Let $\boldsymbol{e}_\alpha$ have a magnitude, or length, such that the displacement vector is given by

$$dx^\alpha \boldsymbol{e}_\alpha,\ \alpha \text{ not summed.}$$

But the magnitude of the displacement is given from the fundamental quadratic differential form, which in this case has the value

$$ds^2 = g_{\alpha\alpha}(dx^\alpha)^2,\ \alpha \text{ not summed.}$$

Hence $\boldsymbol{e}_\alpha$ is such that $\boldsymbol{e}_\alpha \cdot \boldsymbol{e}_\alpha = g_{\alpha\alpha}$, where we use the notation $\boldsymbol{e}_\alpha \cdot \boldsymbol{e}_\alpha$ to denote the scalar product of $\boldsymbol{e}_\alpha$ with itself. Then for an arbitrary infinitesimal displacement, as from M to M', we have

$$\boldsymbol{dM} = dx^\alpha \boldsymbol{e}_\alpha,\ \alpha \text{ summed}$$

and

$$\begin{aligned} ds^2 &= \boldsymbol{dM} \cdot \boldsymbol{dM} = (dx^\alpha \boldsymbol{e}_\alpha) \cdot (dx^\beta \boldsymbol{e}_\beta) \\ &= (\boldsymbol{e}_\alpha \cdot \boldsymbol{e}_\beta) dx^\alpha dx^\beta = g_{\alpha\beta} dx^\alpha dx^\beta. \end{aligned}$$

Since $\boldsymbol{e}_\alpha \cdot \boldsymbol{e}_\beta = g_{\alpha\beta}$, a *necessary and sufficient condition that the parametric lines be orthogonal at every point* is that

$$g_{\alpha\beta} \equiv 0 \text{ for } \alpha \neq \beta.$$

From our conception of a coördinate system, it follows that the vectors $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_N$ thus defined at each point M of the space are linearly independent. Hence they constitute a basis for vectors in the space *at that point* though not necessarily at any other point. Consider the

surface of a sphere as a two-dimensional space. By a vector being in the space is meant a vector which is tangent to the sphere. At a given point vectors $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ tangent, respectively, to the longitude and latitude lines through that point form a basis for any vector in the space at that point, but they will not serve as a basis for vectors in the space at an arbitrary point.

The following example of an N-dimensional geometry may be worth considering. Imagine a machine whose "states" are indicated by N dial readings somewhat like that of an airplane. The "states" of the machine are called "points," and the ordered dial readings corresponding to a given state are called the coördinates of the point. The state for which each dial reading is zero is called the origin. We suppose it is possible to vary the machine in such a fashion that only one dial reading changes at a time while the others remain stationary. The sequence of states which causes only the α dial reading to change is, then, a one-dimensional set of points, which we call the x^α parametric line. We consider now the meaning in this interpretation of the vector $\boldsymbol{e}_\alpha$. Suppose a slight change of state such that only the α dial reading changes, and let its change be denoted by Δx^α. If now Δx^α approaches zero, the other dials remaining stationary, there will be determined a *limiting change of state*. The vector $\boldsymbol{e}_\alpha$ is a symbol which stands for the *change of state* which would produce *unit* change in the α dial reading; that is, would cause its reading to increase by one, *provided it changed at the same rate as initially.*

We consider now the behavior of the base vectors under a transformation of coördinates. Upon the introduction of a second coördinate system, there will be a new set of parametric lines through an arbitrary point M and a new set of base vectors $\bar{\boldsymbol{e}}_1, \bar{\boldsymbol{e}}_2, \ldots, \bar{\boldsymbol{e}}_N$ at that point. Since at the point M the vectors $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_N$ constitute a basis for all vectors in the space at that point, it must be that the vectors $\bar{\boldsymbol{e}}_i$ are expressible as linear combinations of

the $\boldsymbol{e}_\alpha$. We express an infinitesimal displacement $d\boldsymbol{M}$ in terms of each system:

$$d\boldsymbol{M} = dy^i \bar{\boldsymbol{e}}_i = dx^\alpha \boldsymbol{e}_\alpha.$$

Since, however,

$$dy^i = \frac{\partial y^i}{\partial x^\alpha} dx^\alpha,$$

the above relation is equivalent to

$$\bar{\boldsymbol{e}}_i \frac{\partial y^i}{\partial x^\alpha} dx^\alpha = \boldsymbol{e}_\alpha dx^\alpha.$$

But the dx^α are *arbitrary*, and we conclude that

$$\bar{\boldsymbol{e}}_i \frac{\partial y^i}{\partial x^\alpha} = \boldsymbol{e}_\alpha,$$

or

$$(F) \qquad \bar{\boldsymbol{e}}_i = \boldsymbol{e}_\alpha \frac{\partial x^\alpha}{\partial y^i}.$$

This relation is extremely instructive. First we note that if ξ^β is any contravariant vector, then

$$\begin{aligned} \xi^i \bar{\boldsymbol{e}}_i &= \left(\xi^\beta \frac{\partial y^i}{\partial x^\beta}\right) \boldsymbol{e}_\alpha \frac{\partial x^\alpha}{\partial y^i} = \xi^\beta \boldsymbol{e}_\alpha \frac{\partial y^i}{\partial x^\beta} \frac{\partial x^\alpha}{\partial y^i} \\ &= \xi^\beta \boldsymbol{e}_\alpha \delta^\alpha_\beta = \xi^\alpha \boldsymbol{e}_\alpha; \end{aligned}$$

that is, the *bilinear form* $\xi^\alpha \boldsymbol{e}_\alpha$ *is transformed into itself*, by which we mean that the new bilinear form has precisely the same coefficients as the original one. Two linear transformations which have this property with respect to a bilinear form are said to be *contragredient*. Hence the *coefficients* ξ^α (usually called components) of a contravariant vector transform contragrediently to the base vectors $\boldsymbol{e}_\alpha$; this is the reason for the name *contravariant vector*. The coefficients of a covariant vector transform in the same manner as do the base vectors (cogrediently). We note that the *coefficients* ξ^α *of a contravariant vector are simply its coefficients when expressed in terms of the base vectors* $\boldsymbol{e}_\alpha$.

From (F) we also have

$$\bar{g}_{ij} = \bar{e}_i \cdot \bar{e}_j = e_\alpha \cdot e_\beta \frac{\partial x^\alpha}{\partial y^i} \frac{\partial x^\beta}{\partial y^j} = g_{\alpha\beta} \frac{\partial x^\alpha}{\partial y^i} \frac{\partial x^\beta}{\partial y^j}$$

in agreement with (D).

At an arbitrary point M let the reciprocal system of base vectors to $e_1, e_2, \ldots, e_N$ be introduced. Let them be denoted by $e^1, e^2, \ldots, e^N$; they are defined by

$$e^\beta \cdot e_\alpha = \delta^\beta_\alpha.$$

Since the vectors e_α form a basis at M the vectors e^β must be expressible in the form $e^\beta = \mu^{\beta\alpha} e_\alpha$, where the scalar coefficients $\mu^{\beta\alpha}$ are yet to be determined. Forming the scalar product of each side with e_λ gives

$$\delta^\beta_\lambda = g_{\alpha\lambda} \mu^{\beta\alpha}.$$

A known solution of this system of equations is $\mu^{\beta\alpha} = g^{\beta\alpha}$, where

$$g^{\beta\alpha} = \frac{\text{cofactor of } g_{\alpha\beta} \text{ in the determinant } |g_{\alpha\beta}|}{|g_{\alpha\beta}|},$$

and since the determinant $|g_{\alpha\beta}| \neq 0$, the solution is unique. We then have

$$(G) \qquad e^\beta = g^{\alpha\beta} e_\alpha.$$

It now readily follows that

$$e^\beta \cdot e^\gamma = g^{\alpha\beta} e_\alpha \cdot e^\gamma = g^{\alpha\beta} \delta^\gamma_\alpha = g^{\gamma\beta}.$$

Thus we may interpret the quantities $g^{\alpha\beta}$ as the *scalar products of the reciprocal base vectors in pairs,*

$$(H) \qquad e^\alpha \cdot e^\beta = g^{\alpha\beta} = g^{\beta\alpha}.$$

Solving the system (G) for the e_α in terms of e^β, we obtain the unique solution

$$(I) \qquad e_\alpha = g_{\alpha\beta} e^\beta.$$

Forming the scalar product of each member of (I) with e^γ gives

$$(J) \qquad g_{\alpha\beta}g^{\beta\gamma} = \delta^\gamma_\alpha.$$

We are now in position to ascertain how the vectors e^α transform. We obtain this information from (I), knowing how the vectors e_α and the tensor $g_{\alpha\beta}$ transform. The relation (I) holds in any coördinate system. Then

$$\bar{e}_i = \bar{g}_{ij}\bar{e}^j;$$

hence

$$e_\alpha \frac{\partial x^\alpha}{\partial y^i} = g_{\alpha\beta}\frac{\partial x^\alpha}{\partial y^i}\frac{\partial x^\beta}{\partial y^j}\bar{e}^j.$$

Multiply each side of this equation by $\partial y^i/\partial x^\gamma$ and sum for i. Then

$$e_\gamma = g_{\gamma\beta}\frac{\partial x^\beta}{\partial y^j}\bar{e}^j.$$

We now multiply each side of this equation by $g^{\gamma\lambda}$, sum for γ, and make use of (G) and (J). The result is

$$e^\lambda = \bar{e}^j\frac{\partial x^\lambda}{\partial y^j},$$

or

$$(K) \qquad \bar{e}^i = e^\alpha\frac{\partial y^i}{\partial x^\alpha}.$$

If now ξ_β is any covariant vector,

$$\bar{\xi}_i = \xi_\beta\frac{\partial x^\beta}{\partial y^i}$$

and

$$\bar{\xi}_i\bar{e}^i = \xi_\beta e^\alpha\frac{\partial y^i}{\partial x^\alpha}\frac{\partial x^\beta}{\partial y^i} = \xi_\beta e^\alpha\delta^\beta_\alpha = \xi_\alpha e^\alpha.$$

That is, $\xi_\alpha e^\alpha$ is an *invariant bilinear form* and *the coefficients of a covariant vector are* seen to be simply *the coefficients of the vector when it is expressed in terms of the reciprocal base vectors,* $e^1, e^2, \ldots, e^N$.

From (K) we readily obtain the law of transformation of the quantities $g^{\alpha\beta}$. For

$$\bar{g}^{ij} = \bar{e}^i\cdot\bar{e}^j = \left(e^\alpha\frac{\partial y^i}{\partial x^\alpha}\right)\cdot\left(e^\beta\frac{\partial y^j}{\partial x^\beta}\right).$$

Hence

$$(L) \qquad \bar{g}^{ij} = g^{\alpha\beta}\frac{\partial y^i}{\partial x^\alpha}\frac{\partial y^j}{\partial x^\beta}.$$

Thus the quantities $g^{\alpha\beta}$ constitute a contravariant tensor of the second order which is symmetric.

The terms "contravariant" and "covariant" applied to a tensor in reality tell only how the tensor is described, in the same way that "analytic geometry" does not characterize the type of geometry but merely describes the way in which it is being studied. (See §9.5.)

Let the contravariant description of a vector be ξ^α. We shall now obtain the covariant description of the same vector.

By means of (I),

$$\xi^\alpha \boldsymbol{e}_\alpha = \xi^\alpha(g_{\alpha\beta}\boldsymbol{e}^\beta) = (\xi^\alpha g_{\alpha\beta})\boldsymbol{e}^\beta.$$

If now ξ_β is defined by

$$\xi_\beta = \xi^\alpha g_{\alpha\beta},$$

we have

$$\xi^\alpha \boldsymbol{e}_\alpha = \xi_\beta \boldsymbol{e}^\beta.$$

This is what we mean by the covariant description ξ_α of the vector whose contravariant description is ξ^α. We observe that a necessary and sufficient condition that the contravariant and covariant descriptions of an arbitrary vector be the same is that

$$g_{\alpha\beta} = \delta^\alpha_\beta,$$

which in terms of the local base vectors means that they form a unitary orthogonal set at every point. This is the situation which obtains in the case of a rectangular Cartesian coördinate system.

20.8 Algebra of tensors.

The following operations constitute what is known as the *algebra of tensors*. The reader will be able to verify the statements in each case.

(1) If T is a tensor and φ is a scalar, then φT is a tensor of the same type. (Scalar multiplication.)

(2) If T_1 and T_2 are tensors of the *same* type, then $T_1 + T_2$ is a tensor of the same type. (Addition.)

(3) If T_1 and T_2 are any two tensors, then T_1T_2 is a tensor. (Product or multiplication.)

(4) If $\xi^{\alpha\beta}{}_{\gamma}{}^{\lambda}{}_{\mu}$ is a tensor of the type indicated, then a new tensor with two less indices is obtained by summing a contravariant index against a covariant index. For instance, $\xi^{\alpha\beta}{}_{\beta}{}^{\lambda}{}_{\mu}$ is a tensor of the form $\eta^{\alpha\lambda}{}_{\mu}$. This process is called *contraction.*

We illustrate these processes by the following examples:

(1) If the components of T are $\xi^1, \xi^2, \ldots \xi^N$, the components of φT are $\varphi\xi^1, \varphi\xi^2, \ldots, \varphi\xi^N$.

(2) If the components of T_1 and T_2 are, respectively, $\xi^\alpha{}_\beta$ and $\eta^\alpha{}_\beta$, then the components of $T_1 + T_2$ are $(\xi^\alpha{}_\beta + \eta^\alpha{}_\beta)$.

(3) If the components of T_1 and T_2 are $g_{\alpha\beta}$ and ξ^λ, respectively, the components of T_1T_2 are $\eta_{\alpha\beta}{}^\lambda$ where

$$\eta_{\alpha\beta}{}^\lambda = g_{\alpha\beta}\xi^\lambda.$$

(4) With the same tensor as in (3),

$$\eta_{\lambda\beta}{}^\lambda = g_{\lambda\beta}\xi^\lambda = \xi_\beta.$$

Thus to obtain the covariant description of a contravariant vector ξ^α we may form the tensor $g_{\alpha\beta}\xi^\lambda$ and then contract by setting α equal to λ and summing,

$$\xi_\beta = g_{\lambda\beta}\xi^\lambda.$$

Notationally, this operation results in a lowering of the index. More generally, we see that if any tensor is multiplied by $g_{\alpha\beta}$ (or $g^{\alpha\beta}$) and contracted with respect to one of the indices of $g_{\alpha\beta}$ (or $g^{\alpha\beta}$), it will result in a new tensor having one more covariant (contravariant) index than formerly. The process is called *raising* or *lowering of indices* by means of the fundamental tensor $g_{\alpha\beta}$.

It is important to note that each of these operations applied to tensors yields a tensor. Since a tensor is an

invariant, these operations yield new invariants, which means that all of the results obtained in this way have a significance which is *independent of the particular coördinate system* employed, except of course that it must be among the admitted class of coördinate systems. This independence of the coördinate system property is one which we require of any geometric or physical theorem. If a physicist obtains from a given experiment any result of significance it must be independent of the particular apparatus used and of the observer. That is, the result must be susceptible of verification by another observer using a different piece of apparatus (different coördinate system). Tensor analysis is the most powerful tool yet devised for building up the type of invariants needed in differential geometry and physics.

In general, the components of a tensor change under a transformation of coördinates. However, *if every component of a tensor is zero in one coördinate system, this will be true in every coördinate system*. That is, the vanishing of a tensor is itself an *invariant property*. The importance of this fact cannot be overemphasized. In any geometric or physical theory the vanishing of a tensor describes an invariant situation, which may or may not turn out to be important, but which always merits examination.

We have seen that $\xi^\alpha \mathbf{e}_\alpha$ transforms into itself. From the laws of transformation it is seen that the same is true of $\xi^{\alpha\beta}\mathbf{e}_\alpha\mathbf{e}_\beta$ and more generally for any tensor, for example

$$\xi^{\alpha\beta}{}_\gamma{}^\lambda{}_\mu \mathbf{e}_\alpha \mathbf{e}_\beta \mathbf{e}^\gamma \mathbf{e}_\lambda \mathbf{e}^\mu.$$

A tensor of the second order, such as $\xi^{\alpha\beta}\mathbf{e}_\alpha\mathbf{e}_\beta$, is frequently called a *dyad*. See *Gibbs-Wilson* (7); *Lagally* (13).

Exercises

20.14. Let y^1, y^2, y^3 be rectangular Cartesian coördinates, and let x^1, x^2, x^3 be space polar coördinates. Let $F(y) \equiv y^1y^3 - (y^2)^2$ be a scalar invariant. Obtain the gradient of F in each coördinate system and verify that it is a covariant vector. Verify that the tangent vector of the twisted cubic

$$y^1 = t,\ y^2 = t^2,\ y^3 = t^3$$

is a contravariant vector.

20.15. In the restricted relativity theory the interval ds between two near-by point events is given by

$$(ds)^2 = (dx)^2 + (dy)^2 + (dz)^2 - c^2\,(dt)^2,$$

c being a fixed positive constant which is the same in all admitted coördinate systems. Show that ds is an invariant under the group of Lorentz transformations

$$x = \beta\,(\bar{x} - v\bar{t}),\ y = \bar{y},\ z = \bar{z},\ t = \beta\left(\bar{t} - \frac{v\bar{x}}{c^2}\right),$$

where

$$\beta = \left(1 - \frac{v^2}{c^2}\right)^{-\frac{1}{2}},$$

v being the parameter of the transformation, which is a real constant less than c.

20.16. Establish the fact that the Lorentz transformations actually form a *group*.

20.17. Let x, y be rectangular Cartesian coördinates, and let r, θ be polar coördinates in the plane. Let $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ be base vectors determined by the Cartesian coordinates and let $\bar{\boldsymbol{e}}_1$, $\bar{\boldsymbol{e}}_2$ be those determined by the polar coördinates.

(1) Obtain $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ in terms of $\bar{\boldsymbol{e}}_1$, $\bar{\boldsymbol{e}}_2$, and conversely.

(2) Draw to scale the vectors $\boldsymbol{e}_1$, $\boldsymbol{e}_2$; $\bar{\boldsymbol{e}}_1$, $\bar{\boldsymbol{e}}_2$; $\boldsymbol{e}^1$, $\boldsymbol{e}^2$; $\bar{\boldsymbol{e}}^1$, $\bar{\boldsymbol{e}}^2$ at points M and N, whose Cartesian coördinates are (1, 2) and (3, 4) respectively.

(3) If $\boldsymbol{a}$ and $\boldsymbol{b}$ are *unit* vectors having the directions of $\bar{\boldsymbol{e}}_1$ and $\bar{\boldsymbol{e}}_2$, respectively, show that the velocity vector of a moving particle is given by

$$\frac{dr}{dt}\boldsymbol{a} + r\frac{d\theta}{dt}\boldsymbol{b},$$

and that the acceleration vector is given by

$$\left\{\frac{d^2r}{dt^2} - r\left(\frac{d\theta}{dt}\right)^2\right\}\boldsymbol{a} + \left\{r\frac{d^2\theta}{dt^2} + 2\frac{dr}{dt}\frac{d\theta}{dt}\right\}\boldsymbol{b}.$$

20.18. Under the assumption that $g_{\alpha\beta}\xi^\alpha\xi^\beta$ is positive definite, prove the Cauchy-Schwarz inequality

$$(g_{\alpha\beta}\xi^\alpha\eta^\beta)^2 \leqq (g_{\alpha\beta}\xi^\alpha\xi^\beta)(g_{\alpha\beta}\eta^\alpha\eta^\beta),$$

and hence that the angle between two vectors as above defined is real, not complex.

20.19. If a tensor of the second order is such that $\xi^{\alpha\beta} = -\xi^{\beta\alpha}$, it is said to be *alternating* (or skew-symmetric). Show that the property of a tensor being alternating is an invariant property with respect to the group of transformations.

20.20. If ξ^α and η^β are vectors, the alternating tensor

$$\zeta^{\alpha\beta} = \xi^\alpha\eta^\beta - \xi^\beta\eta^\alpha$$

may be regarded as a generalization of the vector cross product of the two vectors; it is frequently called a *bivector*. Establish the following:

(1) A bivector has at most $N(N-1)/2$ distinct non-zero components.

(2) *Only in the case* $N = 3$ can the coëfficients of a bivector be taken as the coefficients of a vector in the same space. In such a case the bivector is called the *vector cross product* of the two given vectors.

(3) The coefficients of a bivector may be interpreted as the areas obtained by parallel projections of the two-dimensional parallelepiped (parallelogram) whose sides are the given vectors upon the coördinate planes. (This refers, of course, to the local affine coördinate system determined by the point at which the vectors are taken, and the base vectors at that point).

(4) The coefficients $\zeta^{\alpha\beta}$ are the determinants formed from the α and β columns of the matrix

$$\begin{pmatrix} \xi^1 & \xi^2 & \cdots & \xi^N \\ \eta^1 & \eta^2 & \cdots & \eta^N \end{pmatrix}.$$

(5) A necessary and sufficient condition that two vectors of the same type be linearly dependent is that their alternating product vanish.

20.21. Let

$$y^i = y^i(x^1, x^2, \cdots, x^N), \ (i = 1, 2, \cdots, N)$$

be a transformation with functional determinant denoted by

$$\left|\frac{\partial y}{\partial x}\right|.$$

(1) If one transformation is followed by another, the resultant of the two is a transformation whose functional determinant is the product of the functional determinants of the two transformations.

(2) If a transformation is followed by its inverse, the product of the functional determinants is equal to 1; that is,

$$\left|\frac{\partial y}{\partial x}\right|\left|\frac{\partial x}{\partial y}\right| = 1.$$

(3) From (2), or otherwise, deduce that

$$\frac{\partial x^\alpha}{\partial y^i} = \frac{\text{cofactor of } \dfrac{\partial y^i}{\partial x^\alpha} \text{ in } \left|\dfrac{\partial y}{\partial x}\right|}{\left|\dfrac{\partial y}{\partial x}\right|}.$$

(4) If g is the determinant $|g_{\alpha\beta}|$, show that $J^2\bar{g} = g$, where J is the functional determinant

$$J = \left|\frac{\partial y}{\partial x}\right|.$$

§21. Covariant Differentiation

We now consider briefly the differential calculus of tensors. Since we no longer necessarily have a system of *constant* base vectors at our disposal, the process of differentiation becomes somewhat more involved than that of merely differentiating the coefficients of a vector.

21.1 Covariant differentiation.

Let ξ^α be a contravariant vector field whose components are at most functions of the coördinates $x^1, x^2, \ldots, x^N$. We write

$$\boldsymbol{\xi} = \xi^\alpha \boldsymbol{e}_\alpha$$

where $\boldsymbol{e}_\alpha$ form the local system of base vectors at the point M under consideration. Corresponding to a displacement specified by $dx^1, dx^2, \ldots, dx^N$, we raise the question as to what we shall mean by "differential $\boldsymbol{\xi}$," which we denote by the symbol $d\boldsymbol{\xi}$. We recognize that, correspond-

ing to the differential changes in the coördinates, there will be a change in the vector arising from two sources: (1) the *coefficients* of the vector will change, since they are functions of the coördinates, and (2) the *base vectors* in terms of which the given vector is expressed will also change. Let $\Delta\boldsymbol{\xi}$ denote the change in the vector $\boldsymbol{\xi}$ corresponding to the changes $\Delta x^1, \Delta x^2, \ldots, \Delta x^N$ in the coördinates. We have

$$\begin{aligned}\Delta\boldsymbol{\xi} &= (\xi^\alpha + \Delta\xi^\alpha)(\boldsymbol{e}_\alpha + \Delta\boldsymbol{e}_\alpha) - \xi^\alpha\boldsymbol{e}_\alpha \\ &= (\Delta\xi^\alpha)\boldsymbol{e}_\alpha + \xi^\alpha(\Delta\boldsymbol{e}_\alpha) + (\Delta\xi^\alpha)(\Delta\boldsymbol{e}_\alpha).\end{aligned}$$

If we indicate the principal parts of the infinitesimals by the usual notation, and note that the last term on the right is a product of infinitesimals in which the other terms are linear, we have

$$d\boldsymbol{\xi} = (d\xi^\alpha)\boldsymbol{e}_\alpha + \xi^\alpha(d\boldsymbol{e}_\alpha).$$

Since the individual ξ^α are scalar functions of the coördinates, we know from the calculus the meaning of $d\xi^\alpha$, namely,

$$d\xi^\alpha = \frac{\partial \xi^\alpha}{\partial x^\beta}dx^\beta.$$

Also we observe that we shall have a meaning for $d\boldsymbol{\xi}$ as soon as we assign a meaning to $d\boldsymbol{e}_\alpha$. Concerning $d\boldsymbol{e}_\alpha$, we make the following assumptions:

(1) $d\boldsymbol{e}_\alpha$ is a *vector* representable in the local base system.
(2) $d\boldsymbol{e}_\alpha$ depends *linearly* upon $dx^1, dx^2, \ldots, dx^N$.

Hence we write

$$d\boldsymbol{e}_\alpha = \Gamma^\lambda{}_{\alpha\beta}dx^\beta\boldsymbol{e}_\lambda,$$

where the $\Gamma^\lambda{}_{\alpha\beta}$ are functions of the coördinates. Then

$$d\boldsymbol{\xi} = \left(\frac{\partial \xi^\alpha}{\partial x^\beta}\boldsymbol{e}_\alpha + \xi^\alpha\Gamma^\lambda{}_{\alpha\beta}\boldsymbol{e}_\lambda\right)dx^\beta$$

or

$$(A) \qquad d\xi = \left(\frac{\partial \xi^\lambda}{\partial x^\beta} + \xi^\alpha \Gamma^\lambda{}_{\alpha\beta}\right) e_\lambda dx^\beta.$$

From

$$\bar{e}_i = e_\alpha \frac{\partial x^\alpha}{\partial y^i}$$

we have

$$d\bar{e}_i = (de_\alpha)\frac{\partial x^\alpha}{\partial y^i} + e_\alpha \frac{\partial^2 x^\alpha}{\partial y^i \partial y^j} dy^j.$$

In the y-coördinate system let

$$d\bar{e}_i = \bar{\Gamma}^k{}_{ij} \bar{e}_k dy^j.$$

Then

$$\bar{\Gamma}^k_{ij} \bar{e}_k dy^j = \Gamma^\lambda_{\alpha\beta} e_\lambda \frac{\partial x^\alpha}{\partial y^i} dx^\beta + e_\alpha \frac{\partial^2 x^\alpha}{\partial y^i \partial y^j} dy^j.$$

or

$$\bar{\Gamma}^k_{ij} e_\lambda \frac{\partial x^\lambda}{\partial y^k} dy^j = \Gamma^\lambda_{\alpha\beta} e_\lambda \frac{\partial x^\alpha}{\partial y^i} \frac{\partial x^\beta}{\partial y^j} dy^j + e_\lambda \frac{\partial^2 x^\lambda}{\partial y^i \partial y^j} dy^j.$$

Since the vectors e_λ are linearly independent, and since we suppose this relation to hold for dy^i arbitrary, it follows that the quantities $\Gamma^k_{ij}(y)$ and $\Gamma^\lambda_{\alpha\beta}(x)$ are related thus:

$$(B) \qquad \bar{\Gamma}^k_{ij} \frac{\partial x^\lambda}{\partial y^k} = \Gamma^\lambda_{\alpha\beta} \frac{\partial x^\alpha}{\partial y^i} \frac{\partial x^\beta}{\partial y^j} + \frac{\partial^2 x^\lambda}{\partial y^i \partial y^j}.$$

It will be observed that the quantities $\Gamma^\lambda_{\alpha\beta}$ do not form a tensor. However, if they are given in one coördinate system these equations specify the corresponding quantities in any other coördinate system and hence they define an invariant. The equations (B) are known as the *Christoffel transformation equations*.

Perhaps the most important choice of the functions $\Gamma^\lambda_{\alpha\beta}$ is one in which these quantities are related to the fundamental covariant tensor $g_{\alpha\beta}$ in a way which we now indicate. From

$$\bar{g}_{ik} = g_{\alpha\gamma} \frac{\partial x^\alpha}{\partial y^i} \frac{\partial x^\gamma}{\partial y^k}$$

we have

$$\frac{\partial \bar{g}_{ik}}{\partial y^j} = \frac{\partial g_{\alpha\gamma}}{\partial x^\beta}\frac{\partial x^\alpha}{\partial y^i}\frac{\partial x^\beta}{\partial y^j}\frac{\partial x^\gamma}{\partial y^k} + g_{\alpha\gamma}\frac{\partial^2 x^\alpha}{\partial y^i \partial y^j}\frac{\partial x^\gamma}{\partial y^k} + g_{\alpha\gamma}\frac{\partial x^\alpha}{\partial y^i}\frac{\partial^2 x^\gamma}{\partial y^k \partial y^j}.$$

We write two similar equations obtained from this one by a cyclic permutation on the indices. We then form the combination

$$\frac{1}{2}\left(\frac{\partial \bar{g}_{ik}}{\partial y^j} + \frac{\partial \bar{g}_{jk}}{\partial y^i} - \frac{\partial \bar{g}_{ij}}{\partial y^k}\right),$$

which we denote by $\bar{\Gamma}_{k;ij}$. Indicating the same combination of the partial derivatives of the $g_{\alpha\beta}$ with respect to the x's by $\Gamma_{\gamma;\alpha\beta}$, it results that

$$\bar{\Gamma}_{k;ij} = \Gamma_{\gamma;\alpha\beta}\frac{\partial x^\gamma}{\partial y^k}\frac{\partial x^\alpha}{\partial y^i}\frac{\partial x^\beta}{\partial y^j} + g_{\alpha\beta}\frac{\partial^2 x^\alpha}{\partial y^i \partial y^j}\frac{\partial x^\beta}{\partial y^k}$$

where in the reduction we have made use of the symmetry of $g_{\alpha\beta}$ and the fact that

$$\frac{\partial^2 x^\beta}{\partial y^i \partial y^j} = \frac{\partial^2 x^\beta}{\partial y^j \partial y^i}, \text{ etc.}$$

The quantities $\Gamma_{\gamma;\alpha\beta}$ are known as the *Christoffel symbols of the first kind* formed from the $g_{\alpha\beta}$. The Christoffel symbols of the *second kind* denoted by $\Gamma^\lambda{}_{\alpha\beta}$ are defined by

$$\Gamma^\lambda{}_{\alpha\beta} = g^{\gamma\lambda}\Gamma_{\gamma;\alpha\beta}.$$

Since

$$\bar{g}^{kl} = g^{\rho\lambda}\frac{\partial y^k}{\partial x^\rho}\frac{\partial y^l}{\partial x^\lambda}$$

we have from the above equation

$$\begin{aligned}
\bar{\Gamma}^l{}_{ij} &= \Gamma_{\gamma;\alpha\beta}g^{\rho\lambda}\frac{\partial y^k}{\partial x^\rho}\frac{\partial y^l}{\partial x^\lambda}\frac{\partial x^\gamma}{\partial y^k}\frac{\partial x^\alpha}{\partial y^i}\frac{\partial x^\beta}{\partial y^j} + g_{\alpha\beta}g^{\rho\lambda}\frac{\partial y^k}{\partial x^\rho}\frac{\partial y^l}{\partial x^\lambda}\frac{\partial^2 x^\alpha}{\partial y^i \partial y^j}\frac{\partial x^\beta}{\partial y^k} \\
&= \Gamma_{\gamma;\alpha\beta}g^{\gamma\lambda}\frac{\partial y^l}{\partial x^\lambda}\frac{\partial x^\alpha}{\partial y^i}\frac{\partial x^\beta}{\partial y^j} + g_{\alpha\beta}g^{\beta\lambda}\frac{\partial y^l}{\partial x^\lambda}\frac{\partial^2 x^\alpha}{\partial y^i \partial y^j}. \\
&= \Gamma^\lambda{}_{\alpha\beta}\frac{\partial y^l}{\partial x^\lambda}\frac{\partial x^\alpha}{\partial y^i}\frac{\partial x^\beta}{\partial y^j} + \frac{\partial^2 x^\lambda}{\partial y^i \partial y^j}\frac{\partial y^l}{\partial x^\lambda},
\end{aligned}$$

or finally,

$$\bar{\Gamma}^k{}_{ij}\frac{\partial x^\lambda}{\partial y^k} = \Gamma^\lambda{}_{\alpha\beta}\frac{\partial x^\alpha}{\partial y^i}\frac{\partial x^\beta}{\partial y^j} + \frac{\partial^2 x^\lambda}{\partial y^i \partial y^j}.$$

Hence the Christoffel symbols of the second kind formed from the fundamental tensor $g_{\alpha\beta}$ satisfy the equations (B). In the sequel we shall understand that this choice of the quantities $\Gamma^{\lambda}{}_{\alpha\beta}$ has been made, that is, that the $\Gamma^{\lambda}{}_{\alpha\beta}$ are the Christoffel symbols of the second kind formed from the fundamental tensor $g_{\alpha\beta}$.

Let a contravariant vector field ξ be expressed in each of two coördinate systems

$$\xi = \bar{\xi}^i \bar{e}_i = \xi^\alpha e_\alpha.$$

Then

$$\begin{aligned} d\xi &= \left(\frac{\partial \bar{\xi}^i}{\partial y^j} + \bar{\xi}^k \bar{\Gamma}^i{}_{kj}\right)\bar{e}_i dy^j \\ &= \left(\frac{\partial \xi^\lambda}{\partial x^\beta} + \xi^\alpha \Gamma^\lambda{}_{\alpha\beta}\right) e_\lambda dx^\beta, \end{aligned}$$

or

$$\left(\frac{\partial \bar{\xi}^i}{\partial y^j} + \bar{\xi}^k \bar{\Gamma}^i{}_{kj}\right) e_\lambda \frac{\partial x^\lambda}{\partial y^i}\frac{\partial y^j}{\partial x^\beta} dx^\beta = \left(\frac{\partial \xi^\lambda}{\partial x^\beta} + \xi^\alpha \Gamma^\lambda{}_{\alpha\beta}\right) e_\lambda dx^\beta.$$

Since the e_λ are linearly independent and the dx^β are arbitrary, it follows that

$$\left(\frac{\partial \bar{\xi}^i}{\partial y^j} + \bar{\xi}^k \bar{\Gamma}^i{}_{kj}\right)\frac{\partial x^\lambda}{\partial y^i}\frac{\partial y^j}{\partial x^\beta} = \left(\frac{\partial \xi^\lambda}{\partial x^\beta} + \xi^\alpha \Gamma^\lambda{}_{\alpha\beta}\right).$$

Hence the quantities

$$\frac{\partial \xi^\lambda}{\partial x^\beta} + \xi^\alpha \Gamma^\lambda{}_{\alpha\beta}$$

constitute a mixed tensor of the second order, which is called the *covariant derivative* of ξ^λ. We shall use the notation $D\xi^\lambda/\partial x^\beta$ to stand for this operation; that is,

$$\frac{D\xi^\lambda}{\partial x^\beta} \equiv \frac{\partial \xi^\lambda}{\partial x^\beta} + \xi^\alpha \Gamma^\lambda{}_{\alpha\beta}$$

We have seen that if ξ^λ is a contravariant vector, then $\varphi\xi^\lambda$ is likewise where φ is a scalar function. Consider then

$$\frac{D\eta^\lambda}{\partial x^\beta}$$

where η^λ is defined by $\eta^\lambda = \varphi\xi^\lambda$:

$$\begin{aligned}
\frac{D\eta^\lambda}{\partial x^\beta} &= \frac{\partial \eta^\lambda}{\partial x^\beta} + \eta^\alpha \Gamma^\lambda{}_{\alpha\beta} \\
&= \frac{\partial(\varphi\xi^\lambda)}{\partial x^\beta} + (\varphi\xi^\alpha)\Gamma^\lambda{}_{\alpha\beta} = \frac{\partial\varphi}{\partial x^\beta}\xi^\lambda + \varphi\frac{\partial\xi^\lambda}{\partial x^\beta} + (\varphi\xi^\alpha)\Gamma^\lambda{}_{\alpha\beta} \\
&= \frac{\partial\varphi}{\partial x^\beta}\xi^\lambda + \varphi\frac{D\xi^\lambda}{\partial x^\beta}.
\end{aligned}$$

If now we define $D\varphi/\partial x^\beta$, where φ is a scalar function, to mean the ordinary partial derivative $\partial\varphi/\partial x^\beta$, we have

$$\frac{D(\varphi\xi^\lambda)}{\partial x^\beta} = \left(\frac{D\varphi}{\partial x^\beta}\right)\xi^\lambda + \varphi\frac{D\xi^\lambda}{\partial x^\beta}. \tag{1}$$

Also it is evident that

$$\frac{D(\xi^\lambda + \eta^\lambda)}{\partial x^\beta} = \frac{D\xi^\lambda}{\partial x^\beta} + \frac{D\eta^\lambda}{\partial x^\beta}. \tag{2}$$

The reciprocal vectors $\boldsymbol{e}^\lambda$ satisfy the equations

$$\boldsymbol{e}_\alpha \cdot \boldsymbol{e}^\lambda = \delta^\lambda{}_\alpha$$

identically in $x^1, x^2, \ldots, x^N$. Let $\Delta x^1, \Delta x^2, \ldots, \Delta x^N$ be infinitesimals, and consider

$$\Delta(\boldsymbol{e}_\alpha \cdot \boldsymbol{e}^\lambda) = \Delta(\delta^\lambda{}_\alpha) = 0,$$

since $\delta^\lambda{}_\alpha$ = constant.

By $\Delta(\boldsymbol{e}_\alpha \cdot \boldsymbol{e}^\lambda)$ we mean

$$\Delta(\boldsymbol{e}_\alpha \cdot \boldsymbol{e}^\lambda) = (\boldsymbol{e}_\alpha + \Delta\boldsymbol{e}_\alpha) \cdot (\boldsymbol{e}^\lambda + \Delta\boldsymbol{e}^\lambda) - \boldsymbol{e}_\alpha \cdot \boldsymbol{e}^\lambda.$$

Hence, neglecting infinitesimals of higher order,

$$(d\boldsymbol{e}_\alpha) \cdot \boldsymbol{e}^\lambda + \boldsymbol{e}_\alpha \cdot (d\boldsymbol{e}^\lambda) = 0.$$

If we suppose for the moment that

$$d\boldsymbol{e}^\lambda = H^\lambda{}_{\alpha\beta}\boldsymbol{e}^\alpha dx^\beta,$$

where the $H^\lambda{}_{\alpha\beta}$ are functions of the coördinates, we have

$$\Gamma^\mu{}_{\alpha\beta}\boldsymbol{e}_\mu \cdot \boldsymbol{e}^\lambda dx^\beta + \boldsymbol{e}_\alpha \cdot H^\lambda{}_{\mu\beta}\boldsymbol{e}^\mu dx^\beta = 0,$$

or

$$\Gamma^\lambda{}_{\alpha\beta}dx^\beta + H^\lambda{}_{\alpha\beta}dx^\beta = 0.$$

Since the dx^β are arbitrary, it follows that

$$H^\lambda{}_{\alpha\beta} = -\Gamma^\lambda{}_{\alpha\beta}.$$

We are now in position to differentiate any tensor. Consider, then,

$$\begin{aligned} d(\xi^\lambda{}_\mu e_\lambda e^\mu) &= (d\xi^\lambda{}_\mu)e_\lambda e^\mu + \xi^\lambda{}_\mu(de_\lambda)e^\mu + \xi^\lambda{}_\mu e_\lambda(de^\mu) \\ &= (d\xi^\lambda{}_\mu)e_\lambda e^\mu + \xi^\lambda{}_\mu \Gamma^\alpha{}_{\lambda\beta} e_\alpha dx^\beta e^\mu - \xi^\lambda{}_\mu e_\lambda \Gamma^\mu{}_{\alpha\beta} e^\alpha dx^\beta \\ &= \left[\frac{\partial \xi^\lambda{}_\mu}{\partial x^\beta} + \xi^\alpha{}_\mu \Gamma^\lambda{}_{\alpha\beta} - \xi^\lambda{}_\alpha \Gamma^\alpha{}_{\mu\beta}\right] e_\lambda e^\mu dx^\beta. \end{aligned}$$

One can show, as in the case of a contravariant vector above, that the quantities in the brackets [ ] constitute a tensor with one contravariant index and two covariant indices. We write

$$\frac{D(\xi^\lambda{}_\mu)}{\partial x^\beta} \equiv \frac{\partial \xi^\lambda{}_\mu}{\partial x^\beta} + \xi^\alpha{}_\mu \Gamma^\lambda{}_{\alpha\beta} - \xi^\lambda{}_\alpha \Gamma^\alpha{}_{\mu\beta},$$

which is called the *covariant derivative of the tensor*. Thus from a given tensor there can be constructed additional tensors by the process of differentiation. Having arrived at the process, we can forget about the base vectors and deal directly with the coefficients of the tensor in forming its covariant derivative.

The following exercises contain some significant results in this connection.

Exercises

21.1. The Christoffel symbols have the following symmetry properties:

$$\Gamma_{\lambda;\alpha\beta} = \Gamma_{\lambda;\beta\alpha}, \text{ and } \Gamma^\lambda{}_{\alpha\beta} = \Gamma^\lambda{}_{\beta\alpha}.$$

21.2. Show that

$$\Gamma_{\gamma;\alpha\beta} + \Gamma_{\alpha;\beta\gamma} = \frac{\partial g_{\alpha\gamma}}{\partial x^\beta}.$$

By means of this, or otherwise, prove the important result: A necessary and sufficient condition that the Christoffel symbols vanish identically is that the $g_{\alpha\beta}$ be constants.

21.3. Any coördinate system in which the $g_{\alpha\beta}$ are constants is called an *affine* coördinate system. Show that in an affine coördinate system covariant differentiation reduces to "ordinary" differentiation.

21.4. Establish the *fundamental result* $\dfrac{D(g_{\alpha\beta})}{\partial x^\gamma} = 0$. This means that in the differentiation process the $g_{\alpha\beta}$ behave as constants.

21.5. (1) If $\zeta^\alpha{}_\beta = \xi^\alpha \eta_\beta$, show that

$$\frac{D\zeta^\alpha{}_\beta}{\partial x^\lambda} = \frac{D\xi^\alpha}{\partial x^\lambda}\eta_\beta + \xi^\alpha \frac{D\eta_\beta}{\partial x^\lambda}.$$

This result together with (1) and (2), developed above, shows that the usual rules for differentiating a sum and a product hold in the case of covariant differentiation.

(2) If φ is a scalar function expressed as a contraction of tensors, its covariant derivative is the same as its ordinary derivative. Prove this theorem, supposing φ to be given by

$$\varphi = \xi^\alpha \eta_\alpha.$$

21.2 Geodesics.

We now consider an important class of curves, known as *geodesics*, which play a role analogous to that of the straight lines in Euclidean geometry. We take the "shortest-distance property" as the defining property of these curves.

Let C be a curve in the space which joins two fixed points A and B. Let the equations of the curve be

$$x^\alpha = x^\alpha(t), \ (\alpha = 1, 2, \cdots, N)$$

and let A correspond to $t = t_1$ and B to $t = t_2$. The arc length of the curve from A to B is then given by

$$\int_{t_1}^{t_2} \sqrt{g_{\alpha\beta}\dot{x}^\alpha \dot{x}^\beta}\, dt,$$

where the dots denote ordinary derivatives with respect to t; that is,

$$\dot{x}^\alpha = \frac{dx^\alpha}{dt}.$$

We now "embed" the curve C in a one-parameter family of comparison curves each of which passes through the fixed points A and B. Letting the coördinates of a point on a curve of the family be denoted by $(z^1, z^2, \ldots, z^N)$, we may take

$$z^\alpha = x^\alpha(t) + \epsilon\varphi^\alpha(t),$$

where $\varphi^\alpha(t)$ satisfy suitable continuity properties and are such that

$$\varphi^\alpha(t_1) = \varphi^\alpha(t_2) = 0,$$

but which are otherwise *arbitrary*, and where ϵ is a parameter (variable) independent of t. We observe that the curve C is a member of the family given by $\epsilon = 0$. For convenience, denote the integrand $\sqrt{g_{\alpha\beta}\dot{z}^\alpha\dot{z}^\beta}$ by $F(z, \dot{z})$. The arc length from A to B of a curve of the family naturally depends upon the particular curve selected; that is, the arc length is a function of ϵ. We write

$$J(\epsilon) = \int_{t_1}^{t_2} F(z, \dot{z})dt = \int_{t_1}^{t_2} F(x^\alpha + \epsilon\varphi^\alpha, \dot{x}^\alpha + \epsilon\dot{\varphi}^\alpha)dt.$$

Suppose the curve C has the shortest arc length among the class of curves considered. This means that $\epsilon = 0$ furnishes a minimum to the function $J(\epsilon)$. By the differential calculus, a necessary condition for this is

$$\left.\frac{dJ(\epsilon)}{d\epsilon}\right]_{\epsilon=0} = 0.$$

Imposing this condition we have

$$0 = \int_{t_1}^{t_2}\left(\frac{\partial F}{\partial x^\alpha}\varphi^\alpha + \frac{\partial F}{\partial \dot{x}^\alpha}\dot{\varphi}^\alpha\right)dt,$$

where now, since ϵ has been set equal to zero, the arguments of F are x and $\dot{x}$ belonging to the curve C. Upon integrating the second term "by parts," one obtains

$$0 = \left.\frac{\partial F}{\partial \dot{x}^\alpha}\varphi^\alpha\right]_{t_1}^{t_2} + \int_{t_1}^{t_2}\varphi^\alpha\left(\frac{\partial F}{\partial x^\alpha} - \frac{d}{dt}\frac{\partial F}{\partial \dot{x}^\alpha}\right)dt.$$

But by hypothesis the φ^α vanish at the end points, and hence

$$\int_{t_1}^{t_2} \varphi^\alpha \left(\frac{d}{dt} \frac{\partial F}{\partial \dot{x}^\alpha} - \frac{\partial F}{\partial x^\alpha} \right) dt = 0.$$

Clearly a sufficient condition for this is that the curve C satisfies the *system of differential equations*

$$(C) \qquad \frac{d}{dt} \frac{\partial F}{\partial \dot{x}^\alpha} - \frac{\partial F}{\partial x^\alpha} = 0, \ (\alpha = 1, 2, \cdots, N).$$

On account of the arbitrariness of the functions φ^α, it can be shown to be also a necessary condition.

Any curve satisfying the system of differential equations (C) is called a *geodesic*. We have just seen that any curve having the "shortest-distance" property must satisfy these equations. However, it must not be concluded that the converse is necessarily true, just as it is not true that every point on a curve $y = f(x)$ at which the slope is zero is necessarily a maximum or minimum point.

Performing the indicated differentiation in (C), one obtains

$$\frac{\frac{d}{dt}(g_{\lambda\beta}\dot{x}^\beta)}{\sqrt{g_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta}} - \frac{g_{\lambda\beta}\dot{x}^\beta}{2\,(g_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta)^{\frac{3}{2}}} \frac{d}{dt}(g_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta) - \frac{\frac{\partial g_{\alpha\beta}}{\partial x^\lambda}\dot{x}^\alpha\dot{x}^\beta}{2\sqrt{g_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta}} = 0.$$

Let now the parameter t be selected as *arc length* s along the curve. Then

$$g_{\alpha\beta}\frac{dx^\alpha}{ds}\frac{dx^\beta}{ds} \overset{s}{\equiv} 1,$$

and the above equations reduce to

$$\frac{d}{ds}\left(g_{\lambda\beta}\frac{dx^\beta}{ds}\right) - \frac{1}{2}\frac{\partial g_{\alpha\beta}}{\partial x^\lambda}\frac{dx^\alpha}{ds}\frac{dx^\beta}{ds} = 0,$$

or

$$g_{\lambda\beta}\frac{d^2x^\beta}{ds^2} + \frac{\partial g_{\lambda\beta}}{\partial x^\alpha}\frac{dx^\alpha}{ds}\frac{dx^\beta}{ds} - \frac{1}{2}\frac{\partial g_{\alpha\beta}}{\partial x^\lambda}\frac{dx^\alpha}{ds}\frac{dx^\beta}{ds} = 0,$$

or

$$g_{\lambda\beta}\frac{d^2x^\beta}{ds^2} + \Gamma_{\lambda;\alpha\beta}\frac{dx^\alpha}{ds}\frac{dx^\beta}{ds} = 0,$$

or finally

$$(D) \qquad \frac{d^2x^\lambda}{ds^2} + \Gamma^\lambda_{\alpha\beta}\frac{dx^\alpha}{ds}\frac{dx^\beta}{ds} = 0, \ (\lambda = 1, 2, \cdots, N).$$

These are, then, the equations which must be satisfied by every geodesic when referred to its arc length as parameter.

Exercises

21.6. The left members of equations (D) can be written in the form $\dfrac{D\left(\dfrac{dx^\lambda}{ds}\right)}{ds}$ and hence constitute a *contravariant vector.*

21.7. In an affine coördinate system a geodesic referred to its arc length satisfies the system of differential equations

$$\frac{d^2x^\lambda}{ds^2} = 0$$

and hence the coördinates of a point on the geodesic are linear functions of s.

21.3 Equations of parallelism.

Let C be a curve

$$x^\alpha = x^\alpha(t)$$

and let ξ be a contravariant vector field defined at points of C. Consider the system of linear, homogeneous, differential equations

$$(E) \qquad \frac{d\xi^\lambda}{dt} + \Gamma^\lambda_{\alpha\beta}\xi^\alpha\frac{dx^\beta}{dt} = 0, \ (\lambda = 1, 2, \cdots N).$$

which may be written in the form

$$\frac{D\xi^\lambda}{dt} = 0, \ (\lambda = 1, 2, \cdots, N).$$

Existence theorems tell us that there exist N linearly independent vectors $\xi^\lambda_{(1)}, \xi^\lambda_{(2)}, \ldots, \xi^\lambda_{(N)}$ satisfying these equations and being such that any solution ξ^λ is expressible as a linear combination of them with constant coefficients.

Let ξ^λ and η^λ be any two solutions of the system (E). We now show that $g_{\alpha\beta}\xi^\alpha\eta^\beta$ = constant along the curve. This is equivalent to

$$\frac{d}{dt}(g_{\alpha\beta}\xi^\alpha\eta^\beta) \equiv 0.$$

Since $g_{\alpha\beta}\xi^\alpha\eta^\beta$ is a scalar, by Exercises **21.5,** part (2), and **21.4,** we have

$$\frac{d}{dt}(g_{\alpha\beta}\xi^\alpha\eta^\beta) = \frac{D(g_{\alpha\beta}\xi^\alpha\eta^\beta)}{dt} = g_{\alpha\beta}\left(\frac{D\xi^\alpha}{dt}\right)\eta^\beta + g_{\alpha\beta}\xi^\alpha\left(\frac{D\eta^\beta}{dt}\right) = 0.$$

Therefore $g_{\alpha\beta}\xi^\alpha\eta^\beta$ = constant along the curve. Similarly $g_{\alpha\beta}\xi^\alpha\xi^\beta$ = constant, and hence the angle θ between the vectors $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ remains constant. Thus any two vectors satisfying the equations (E) have the properties that their lengths remain constant, and the angle θ remains constant along the curve, θ being the angle between the vectors. For these reasons the equations (E) are said to define a *parallel displacement* of a contravariant vector along the curve.

Exercises

21.8. If C is a curve referred to its arc length as parameter, a necessary and sufficient condition that its unit tangent vector remain parallel to itself along the curve is that C be a geodesic. Thus in a sense the "straightness" property of straight lines is carried over to geodesics.

21.9. When referred to an affine coördinate system, a necessary and sufficient condition that ξ^α be a parallel vector field along an arbitrary curve is that ξ^α be a constant vector field; that is, that the coefficients ξ^α be constants.

21.10. In the case of an affine coördinate system, the local base vectors $\boldsymbol{e}_\alpha$ satisfy the equations of parallelism along an arbitrary path. This means that the local system of base vectors

at a point M' may be regarded as the system e_α at an arbitrary point M which has undergone a parallel displacement to the point M'.

21.4 Curved spaces.

If it is possible to introduce into the geometry, determined by a coördinate system x and a fundamental quadratic differential form $g_{\alpha\beta}dx^\alpha dx^\beta$, an affine coördinate system the space is said to be *flat* and the geometry is called *Euclidean*. If this is impossible, the space is said to be *curved* and the geometry is said to be *Riemannian*.

First we remark that, if an affine coördinate system exists, then there exists a rectangular Cartesian coördinate system, for all we need to do to bring this about is to norm and orthogonalize the constant base vectors.

From the definition of an affine coördinate system, the results of Exercise **21.2** and the Christoffel transformations, a necessary and sufficient condition that it be possible to introduce an affine coördinate system by means of a transformation of coördinates is that the system of differential equations

$$\delta_{ij}\frac{\partial y^i}{\partial x^\alpha}\frac{\partial y^j}{\partial x^\beta} = g_{\alpha\beta} \text{ where } \delta_{ij} = \begin{cases} 1 \text{ for } i = j \\ 0 \text{ for } i \neq j \end{cases}.$$

$$\frac{\partial^2 y^i}{\partial x^\alpha \partial x^\beta} = \Gamma^\lambda{}_{\alpha\beta}\frac{\partial y^i}{\partial x^\lambda}$$

shall possess a solution

$$y^i = y^i(x^1, x^2, \cdots, x^N), \ (i = 1, 2, \cdots, N).$$

This problem is too complicated for us profitably to undertake here (cf. *Veblen* (50), Chapter V). It is shown, however, that a necessary and sufficient condition that the system of equations possess a solution is that the *curvature tensor*

$$B^\lambda{}_{\alpha\beta\gamma} \equiv \frac{\partial \Gamma^\lambda_{\alpha\beta}}{\partial x^\gamma} - \frac{\partial \Gamma^\lambda_{\alpha\gamma}}{\partial x^\beta} + \Gamma^\rho_{\alpha\beta}\Gamma^\lambda_{\rho\gamma} - \Gamma^\rho_{\alpha\gamma}\Gamma^\lambda_{\rho\beta}$$

vanish identically. According to this result we can tell in any given instance whether or not there exists a transforma-

tion leading to a Cartesian coördinate system. Moreover, the application of the test involves only ordinary partial differentiation combined with simple algebraic processes which can always be carried out.

Exercises

21.11. The fundamental form for Euclidean geometry referred to plane polar coördinates is

$$ds^2 = (dx^1)^2 + (x^1)^2(dx^2)^2.$$

Verify that the curvature tensor vanishes identically and, hence, that a rectangular Cartesian coördinate system can be introduced.

21.12. The linear element on the sphere of radius 1 referred to co-latitude x^1 and longitude x^2 was found to be

$$ds^2 = (dx^1)^2 + \sin^2 x^1 (dx^2)^2.$$

Verify that the curvature tensor in this case does not vanish and, hence, that there does not exist a Cartesian coördinate system on the sphere.

21.13. If φ is a scalar function, verify that

$$\frac{D^2\varphi}{\partial x^\alpha \partial x^\beta} = \frac{D^2\varphi}{\partial x^\beta \partial x^\alpha}.$$

However, if ξ^λ is an arbitrary vector, a necessary and sufficient condition that

$$\frac{D^2\xi^\lambda}{\partial x^\alpha \partial x^\beta} = \frac{D^2\xi^\lambda}{\partial x^\beta \partial x^\alpha}$$

is that the curvature tensor vanish.

21.5 Divergence, curl, Laplacian.

We have seen that if $\varphi(x)$ is any scalar invariant function, $\partial\varphi/\partial x^\alpha$ is a covariant vector which is called the (tensor) *gradient* of $\varphi(x)$. We now consider the forms which the divergence, curl, and Laplacian assume in a general coördinate system.

Let ξ^α be a contravariant vector field. Then

$$\frac{1}{\sqrt{g}} \frac{\partial}{\partial x^\alpha}(\sqrt{g}\,\xi^\alpha)$$

is a scalar invariant which is known as the *divergence* of the vector field. We shall prove directly that the given expression is actually a scalar invariant. Consider then

$$\frac{1}{\sqrt{\bar{g}}}\frac{\partial}{\partial y^i}(\sqrt{\bar{g}}\xi^i) = \frac{1}{2\bar{g}}\frac{\partial \bar{g}}{\partial y^i}\xi^i + \frac{\partial \xi^i}{\partial y^i}$$
$$= \frac{1}{2}\frac{\partial}{\partial y^i}(\log \bar{g})\xi^i + \frac{\partial \xi^i}{\partial y^i}.$$

From Exercise **20.21**, $\log \bar{g} = \log g - 2 \log J$, where J is the functional determinant

$$J = \left|\frac{\partial y}{\partial x}\right|.$$

Hence

$$\frac{1}{\sqrt{\bar{g}}}\frac{\partial}{\partial y^i}(\sqrt{\bar{g}}\xi^i) = \frac{1}{2}\frac{\partial}{\partial x^\alpha}[\log g - 2 \log J]\xi^i\frac{\partial x^\alpha}{\partial y^i} + \frac{\partial \xi^i}{\partial y^i}$$
$$= \frac{1}{2}\frac{\partial}{\partial x^\alpha}(\log g)\xi^\alpha + \frac{\partial \xi^\alpha}{\partial x^\alpha} + \left\{-\frac{\partial(\log J)}{\partial x^\alpha} + \frac{\partial^2 y^i}{\partial x^\alpha \partial x^\beta}\frac{\partial x^\beta}{\partial y^i}\right\}\xi^\alpha.$$

One can show that the quantity in the braces { } vanishes. From the rule for differentiating a determinant we have

$$\frac{\partial}{\partial x^\alpha}\left|\frac{\partial y}{\partial x}\right| = \frac{\partial^2 y^i}{\partial x^\beta \partial x^\alpha}A_i^\beta,$$

where A_i^β is the cofactor of $\dfrac{\partial y^i}{\partial x^\beta}$ in $\left|\dfrac{\partial y}{\partial x}\right|$.

But by Exercise **20.21**

$$A_i^\beta = \frac{\partial x^\beta}{\partial y^i}\left|\frac{\partial y}{\partial x}\right|.$$

Hence

$$\frac{\partial\left|\frac{\partial y}{\partial x}\right|}{\partial x^\alpha} \Big/ \left|\frac{\partial y}{\partial x}\right| = \frac{\partial}{\partial x^\alpha}(\log J) = \frac{\partial^2 y^i}{\partial x^\beta \partial x^\alpha}\frac{\partial x^\beta}{\partial y^i},$$

and therefore

$$\frac{1}{\sqrt{\bar{g}}}\frac{\partial}{\partial y^i}(\sqrt{\bar{g}}\bar{\xi}^i) = \frac{1}{\sqrt{g}}\frac{\partial}{\partial x^\alpha}(\sqrt{g}\xi^\alpha),$$

which is the result we wished to establish.

Let ξ_α be a covariant vector field. From the law of transformation

$$\bar{\xi}_i = \xi_\alpha \frac{\partial x^\alpha}{\partial y^i}$$

we obtain by ordinary differentiation

$$\frac{\partial \bar{\xi}_i}{\partial y^j} = \frac{\partial \xi_\alpha}{\partial x^\beta}\frac{\partial x^\alpha}{\partial y^i}\frac{\partial x^\beta}{\partial y^j} + \xi_\alpha \frac{\partial^2 x^\alpha}{\partial y^i \partial y^j}.$$

Hence, when we use

$$\frac{\partial^2 x^\alpha}{\partial y^i \partial y^j} = \frac{\partial^2 x^\alpha}{\partial y^j \partial y^i}$$

it follows that

$$\left(\frac{\partial \bar{\xi}_i}{\partial y^j} - \frac{\partial \bar{\xi}_j}{\partial y^i}\right) = \left(\frac{\partial \xi_\alpha}{\partial x^\beta} - \frac{\partial \xi_\beta}{\partial x^\alpha}\right)\frac{\partial x^\alpha}{\partial y^i}\frac{\partial x^\beta}{\partial y^j}.$$

Therefore

$$\xi_{\alpha\beta} \equiv \frac{\partial \xi_\alpha}{\partial x^\beta} - \frac{\partial \xi_\beta}{\partial x^\alpha}$$

constitute an alternating covariant tensor of rank two. This tensor is a generalization of the *curl* of a vector field in a three-dimensional space.

In the case of three dimensions, an alternating covariant tensor of the second order possesses at most three distinct non-zero coefficients. Given, then, any alternating covariant tensor of the second order $\xi_{\alpha\beta}$, there is determined by the invariant form $\xi_{\alpha\beta}e^\alpha e^\beta$ a unique contravariant vector

$$\xi_{12}(e^1 \times e^2) + \xi_{23}(e^2 \times e^3) + \xi_{31}(e^3 \times e^1),$$

which we write in the form

$$\xi_{12}\frac{e_3}{[e_1e_2e_3]} + \xi_{23}\frac{e_1}{[e_1e_2e_3]} + \xi_{31}\frac{e_2}{[e_1e_2e_3]}.$$

We have seen (Exercise **9.7**) that $[e_1e_2e_3] = \sqrt{g}$, where g denotes the determinant $|g_{\alpha\beta}|$, and hence we may write the above vector in the form

$$\frac{\xi_{23}}{-\sqrt{g}}e_1 + \frac{\xi_{31}}{-\sqrt{g}}e_2 + \frac{\xi_{12}}{-\sqrt{g}}e_3,$$

where we have used the negative square root of g in order to give the sign customarily used in the curl. Now then if ξ_α is any *covariant vector field, its curl is the contravariant vector*

$$\frac{1}{-\sqrt{g}}\left\{\left(\frac{\partial\xi_2}{\partial x^3} - \frac{\partial\xi_3}{\partial x^2}\right)e_1 + \left(\frac{\partial\xi_3}{\partial x^1} - \frac{\partial\xi_1}{\partial x^3}\right)e_2 + \left(\frac{\partial\xi_1}{\partial x^2} - \frac{\partial\xi_2}{\partial x^1}\right)e_3\right\}.$$

Exercises

21.14. Obtain the expression for the divergence of a contravariant vector field in terms of space polar coördinates.

21.15. Show that if ξ_α is a covariant vector field, then

$$\frac{1}{\sqrt{g}}\frac{\partial}{\partial x^\alpha}(\sqrt{g}g^{\alpha\beta}\xi_\beta)$$

is a scalar invariant. It is called the divergence of the covariant vector field.

21.16. If in the preceding exercise the covariant vector ξ_β is the gradient of a scalar field—that is, if $\xi_\beta = \partial\varphi/\partial x^\beta$—its divergence is known as the Laplacian of φ. We write $\nabla^2\varphi$ for the Laplacian of φ. Obtain $\nabla^2\varphi$ in space polar coördinates.

21.17. If G denotes the determinant $|g^{\alpha\beta}|$, show that $G = 1/g$, and hence, or otherwise, that $[e^1e^2e^3] = \sqrt{G}$. (The dimensionality $N = 3$ is implied by the notation.)

21.18. If ξ_α is the gradient of a scalar field, its curl vanishes identically.

References

(1) Bouligand, Georges, *Leçons de géométrie vectorielle.* Paris: Vuibert, 1924.

(2) Bouligand, G., and Rabate, G., *Initiation aux méthodes vectorielles et aux applications géométriques de l'analyse.* Paris: Vuibert, 1926.

(3) Bricard, Raoul, *Le calcul vectoriel.* Paris: Colin, 1929.

(4) Burali-Forti, C., and Marcolongo, R., *Analisi vettoriale generale e applicazioni.* Volume Primo. *Trasformazioni lineari.* Bologna: Zanichelli, 1929.

(5) Coffin, Joseph George, *Vector Analysis.* New York: Wiley, 1911.

(6) Gans, Richard, *Vector Analysis with Applications to Physics.* Tr. by Winifred M. Deans. London: Blackie, 1932.

(7) Gibbs, J. Willard, and Wilson, Edwin Bidwell, *Vector Analysis.* New Haven: Yale Univerity Press, 1901.

(8) Grassmann, Hermann, *Die lineale Ausdehnungslehre.* Leipzig: Wigand, 1878.

(9) Haas, Arthur, *Vektoranalysis.* Berlin u. Leipzig: de Gruyter, 1929.

(10) Ignatowsky, W. v., *Die Vektoranalysis.* Teil I. Leipzig u. Berlin: Teubner, 1926.

(11) Juvet, Gustave, *Leçons d'analyse vectorielle.* Première partie. *Géométrie différentielle des courbes et des surfaces. Théorie mathématique des champs.* Paris: Gauthier-Villars, 1933.

(12) ——— *Leçons d'analyse vectorielle.* Deuxième partie. *Applications de l'analyse vectorielle. Introduction à la physique mathématique.* Paris: Gauthier-Villars, 1935.

(13) Lagally, Max, *Vorlesungen über Vektor-Rechnung.* Leipzig: Akademische Verlagsgesellschaft M.B.H., 1934.

(14) Lotze, Alfred, *Punkt- und Vektor-Rechnung.* Berlin u. Leipzig: de Gruyter, 1929.

(15) Phillips, H. B., *Vector Analysis.* New York: Wiley, 1933.

(16) Ramos, T. A., *Leçons sur le calcul vectoriel.* Paris: Blanchard, 1930.

(17) Runge, C., *Vector Analysis.* Tr. by H. Levy. New York: Dutton, 1919.

(18) Shorter, L. R., *Introduction to Vector Analysis.* London: MacMillan, 1931.

(19) Spielrein, Jean, *Lehrbuch der Vektorrechnung.* Stuttgart: Wittwer, 1926.

(20) Valentiner, Siegfried, *Vektoranalysis.* Leipzig: de Gruyter, 1929.

(21) Weatherburn, C. E., *Elementary Vector Analysis with Application to Geometry and Physics.* London: Bell, 1928.

(22) ——— *Advanced Vector Analysis with Application to Mathematical Physics.* London: Bell, 1928.

(23) Wills, A. P., *Vector Analysis with an Introduction to Tensor Analysis.* New York: Prentice-Hall, 1931.
Craig, Homer V., *Vector and Tensor Analysis,* New York: McGraw-Hill, 1943.

(24) Ames, Joseph Sweetman, and Murnaghan, Francis D., *Theoretical Mechanics.* Boston: Ginn, 1929.

(25) Beck, H., *Einführung in die Axiomatik der Algebra.* Leipzig u. Berlin: de Gruyter, 1926.

(26) Blaschke, Wilhelm, *Vorlesungen über Differentialgeometrie.* I. Berlin: Springer, 1930.

(27) Bôcher, Maxime, *Introduction to Higher Algebra.* New York: Macmillan, 1915.

(28) Cartan, E., *Leçons sur la géométrie des espaces de Riemann.* Paris: Gauthier-Villars, 1928.

(29) Carathéodory, Constantin, *Vorlesungen über reelle Funktionen.* Leipzig u. Berlin: Teubner, 1918.

(30) Dickson, Leonard Eugene, *Algebras and Their Arithmetics.* Chicago: University of Chicago Press, 1923.

(31) Eddington, A. S., *The Mathematical Theory of Relativity.* Cambridge: Cambridge University Press, 1924.

(32) Eisenhart, Luther Pfahler, *Riemannian Geometry.* Princeton: Princeton University Press, 1926.

(33) Gibson, George A., *Advanced Calculus.* London: MacMillan, 1931.

(34) Goursat, Edouard, *A Course in Mathematical Analysis.* Vol. 1. Tr. by Earle Raymond Hedrick. Boston: Ginn, 1904.

(35) Graustein, William C., *Introduction to Higher Geometry.* New York: Macmillan, 1930.

(36) Hardy, A. S., *Elements of Quaternions.* Boston: Ginn-Heath, 1881.

(37) Jeans, J. H., *An Elementary Treatise on Theoretical Mechanics.* Boston: Ginn, 1907.

(38) Juvet, G., *Introduction au calcul tensoriel et au calcul différentiel absolu.* Paris: Blanchard, 1922.
(39) Kelland, P., and Tait, P. G., *Introduction to Quaternions.* London: MacMillan, 1882.
(40) Kellogg, Oliver Dimon, *Foundations of Potential Theory.* Berlin: Springer, 1929.
(41) Kowalewski, Gerhard, *Einführung in die Determinantentheorie.* Berlin u. Leipzig: de Gruyter, 1925.
(42) Love, A. E. H., *Theoretical Mechanics.* London: MacMillan, 1921.
(43) Mach, Ernst, *The Science of Mechanics.* Tr. by Thomas J. McCormick. Chicago: Open Court, 1907.
(44) Murnaghan, Francis D., *Vector Analysis and the Theory of Relativity.* Baltimore: The Johns Hopkins Press, 1922.
(45) Osgood, William F., *Advanced Calculus.* New York: Macmillan, 1928.
(46) Osgood, William F., and Graustein, William C., *Plane and Solid Analytic Geometry.* New York: Macmillan, 1927.
(47) Picard, Émile, *Traité d'analyse.* Tome I. Paris: Gauthier-Villars, 1922.
(48) Reynolds, Joseph B., and Weida, Frank M., *Analytic Geometry and the Elements of Calculus.* New York: Prentice-Hall, 1930.
(49) Schreier, O., and Sperner, E., *Einführung in die analytische Geometrie und Algebra.* Band I, II. Leipzig u. Berlin: Teubner, 1931, 1935.
(50) Veblen, Oswald, *Invariants of Quadratic Differential Forms.* Cambridge: Cambridge University Press, 1927.
(51) Veblen, Oswald, and Young, John Wesley, *Projective Geometry.* Vols. I, II. Boston: Ginn, 1916, 1918.
(52) Webster, Arthur Gordon, *Dynamics of Particles and of Rigid, Elastic, and Fluid Bodies.* New York: G. E. Stechert, 1922.
(53) Weyl, Hermann, *Space, Time, Matter.* Tr. by Henry L. Brose. New York: Dutton, 1921.
(54) Woods, Frederick S., *Advanced Calculus.* Boston: Ginn, 1932.
(55) Young, John Wesley, *Lectures on Fundamental Concepts of Algebra and Geometry.* New York: Macmillan, 1925.
Thomas, T. Y., *The Elementary Theory of Tensors with Applications to Geometry and Mechanics.* New York: McGraw-Hill, 1931.

Index